MASTER EN SCIENCES MATHÉMATIQUES

Méthodes d'optimisation non différentiable à évaluations inexactes

Goblet, Jordan

Award date:
2003

FUNDP
Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B-5000 Namur Belgique

Méthodes d'optimisation non différentiable à évaluations inexactes

Mémoire présenté pour l'obtention
du grade de
Licencié en Sciences Mathématiques
par

Jordan Goblet

Promoteur :Jean-Jacques Strodiot

Année Académique 2002-2003

Je remercie spécialement le Professeur Jean-Jacques Strodiot pour sa grande disponibilité et les précieux conseils qu'il m'a apportés tout au long de l'élaboration de ce mémoire.

Je tiens également à remercier mes parents, Aline et mes soeurs pour leur soutien durant ces quatre années d'étude.

Résumé

Nous nous intéressons aux méthodes de minimisation de fonctions convexes non nécessairement différentiables. Nous commençons par introduire les définitions et notions classiques de l'analyse convexe non différentiable. Nous décrivons ensuite une famille de techniques appelées méthodes faisceaux. Les méthodes faisceaux sont alors réinterprétées dans le cadre de travail plus général des méthodes du point proximal. La théorie de convergence de la méthode du point proximal est alors étendue au traitement du cas où la fonction est évaluée inexactement et la méthode du point proximal inexact est présentée. Nous terminons par l'étude de la méthode de Solodov qui fournit une analyse de stabilité pour les méthodes faisceaux standards.

Abstract

We are interested in methods for minimizing nonsmooth convex functions. The paper opens with some classical definitions and notions from nonsmooth convex analysis, and proceeds to provide background on existing solution methods. We introduce a family of techniques called bundle methods. Bundle methods are then reinterpreted in the more general framework of proximal point methods. The convergence theory of the proximal point method is extended to handle inexact function evaluations and the inexact proximal point method is presented. We end by studying the Solodov's method which provides a stability analysis of standard bundle methods.

Table des matières

Chapitre 1

Introduction

Ce mémoire consiste en l'étude de méthodes permettant de résoudre le problème de programmation mathématique sans contrainte

$$\min_{x \in \mathbb{R}^n} \{f(x) \mid x \in \mathbb{R}^n\}$$

où $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe non nécessairement différentiable à valeurs finies. Les principes généraux de l'analyse convexe non différentiable sont succintement décrits dans le chapitre 2 et illustrés par les méthodes de plus forte pente, d'ε plus forte pente et du sous-gradient. Le chapitre 3 introduit les points de vue primal et dual d'une famille de techniques appelées méthodes faisceaux. Les méthodes faisceaux sont réinterprétées dans le cadre de travail plus général des méthodes du point proximal au chapitre 4. La théorie de convergence de la méthode du point proximal est modifiée dans le chapitre 5 afin de permettre des évaluations inexactes de la fonction et de ses sous-gradients. La méthode du point proximal inexact qui gère automatiquement le degré d'approximation est présentée à la fin de ce chapitre. Nous terminons au chapitre 6 par une étude approfondie de la méthode de Solodov qui fournit une analyse de stabilité pour les méthodes faisceaux standards. Les apports récents de la recherche présentés dans ce mémoire sont :

. une présentation unifiée de la théorie de convergence de la méthode du point proximal approximé ;

. une extension de cette théorie qui produit un algorithme (et une analyse de convergence associée) basé sur des évaluations inexactes de la fonction ;

. la méthode de Solodov qui fournit une analyse de stabilité aux méthodes faisceaux.

1.1 Notations

Commençons par définir différentes notations standards. Pour $x, y \in \mathbb{R}^n$, le produit scalaire est noté $x^T y$ ou $< x, y >$. La norme par défaut $\|x\|$ est la norme euclidienne classique $\sqrt{< x, x >}$.
Soit S^m l'espace vectoriel des matrices réelles symétriques de dimension $m \times m$. Pour $X \in S^m$ arbitraire, les valeurs propres de X sont ordonnées et notées $\lambda_1(X) \geq \lambda_2(X) \geq ... \geq \lambda_m(X)$. Une matrice symétrique X est définie positive (négative) si toutes ses valeurs propres sont strictement positives et on écrit $X > 0$ ($X < 0$).
Etant donnée $M \in S^m$ où $M > 0$, nous écrirons $\|x\|_M := \sqrt{< x, Mx >}$ et dans ce contexte, M est appelée métrique.

Chapitre 2

Analyse Convexe Non Différentiable

Puisque ce mémoire consiste en l'étude de méthodes permettant de minimiser des fonctions convexes non différentiables, ce chapitre résume les principes généraux de l'analyse convexe non différentiable.
Afin d'illustrer l'utilité des différents outils développés, les méthodes de plus forte pente, d' ε plus forte pente et du sous-gradient sont succintement décrites. Nous ne développerons pas ces différentes méthodes aux multiples inconvénients, nous préférerons par la suite étudier les méthodes faisceaux pour des raisons à la fois pratiques et d'efficacité.
Le détail des concepts et les preuves des théorèmes sont repris dans [4, 5] et [8].

2.1 Sous-gradient et sous-différentiel

Tout au long de ce mémoire, on considère une fonction $f : \mathbb{R}^n \to \mathbb{R}$ convexe et à valeurs finies. Par conséquent, f est localement Lipschitzienne, continue sur $\mathbb{R}^n$ et différentiable presque partout (par le théorème de Rademacher).
Il existe différentes façons de généraliser la notion de gradient et la plupart d'entre elles sont équivalentes dans le cas de fonctions convexes. En ce qui nous concerne, nous considérons la définition géométrique suivante :

Définition 2.1.1 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe et $x \in \mathbb{R}^n$. Alors, $\xi \in \mathbb{R}^n$ est un sous-gradient de f en x si*

$$f(y) \geq f(x) + < \xi, y - x > \quad \forall y \in \mathbb{R}^n.$$

L'ensemble des sous-gradients de f en x est appelé sous-différentiel de f en x et est noté $\partial f(x)$.

Autrement dit, ξ est la pente d'une fonction affine qui minore f et qui passe par le point $(x, f(x))$. Ce concept généralise le point de vue géométrique du gradient $\nabla f(x)$ qui consiste à définir un plan tangent au graphe de f en $(x, f(x))$.

Exemple 2.1.1 *Considérons la fonction $f(x) = |x|$. Après observation du*

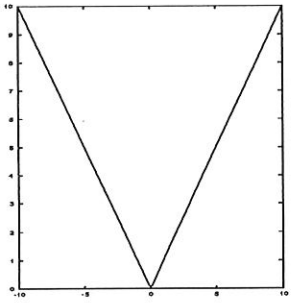

FIG. 2.1 – Graphe de $f(x) = |x|$.

graphe de cette fonction, on peut immédiatement conclure que

$$\partial f(0) = [-1, 1],\ \partial f(x_0) = \{1\} \text{ si } x_0 > 0 \text{ et } \partial f(x_0) = \{-1\} \text{ si } x_0 < 0.$$

Le sous-différentiel peut aussi être caractérisé à l'aide de la dérivée directionnelle.

Proposition 2.1.1 *La propriété suivante est vérifiée :*

$$\xi \in \partial f(x) \iff < \xi, d > \leq f'(x, d) \quad \forall d \in \mathbb{R}^n.$$

Autrement dit, nous avons l'égalité

$$f'(x,d) = \sup_{\xi \in \partial f(x)} < \xi, d > .$$

La proposition suivante fait le lien entre différentiabilité et sous-différentiabilité.

Proposition 2.1.2 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ convexe et $x_0 \in \mathbb{R}^n$. Alors,*

1. *Si f est différentiable en x_0, alors $\partial f(x_0) = \{\nabla f(x_0)\}$.*
2. *Si $\partial f(x_0) = \{\xi\}$ alors f est différentiable en x_0 et $\nabla f(x_0) = \xi$.*

En fait, $\partial f(x)$ est réduit au singleton $\{\nabla f(x)\}$ si et seulement si f est différentiable en x. L'exemple 2.1.1 ne faillit pas à la règle puisque

$$\partial f(x_0) = \{1\} = \{\nabla f(x_0)\} \text{ si } x_0 > 0,$$
$$\partial f(x_0) = \{-1\} = \{\nabla f(x_0)\} \text{ si } x_0 < 0.$$

De plus, le sous-différentiel est un ensemble non vide, convexe et compact. La condition d'optimalité est présentée dans la proposition suivante, elle généralise la propriété familière disant que $\nabla f(x) = 0$ en un point optimal x.

Proposition 2.1.3 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ convexe et $x^* \in \mathbb{R}^n$. Alors, $f(x^*) = \min_x f(x)$ si et seulement si $0 \in \partial f(x^*)$.*

Il existe plusieurs règles de calcul concernant le sous-différentiel qui nous permettent de dériver le sous-différentiel d'une fonction convexe décrite à partir d'autres fonctions convexes. Voici quelques unes des plus simples d'entre elles.

Proposition 2.1.4 *Soient $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ fonctions convexes et soient $\alpha_1, \alpha_2 > 0$. On définit $f : \mathbb{R}^n \to \mathbb{R}$ par*

$$f(x) := \alpha_1 f_1(x) + \alpha_2 f_2(x).$$

Alors,

$$\partial f(x) = \alpha_1 \partial f_1(x) + \alpha_2 \partial f_2(x).$$

Proposition 2.1.5 *Soient $f_1, ..., f_m : \mathbb{R}^n \to \mathbb{R}$ fonctions convexes. On définit $f : \mathbb{R}^n \to \mathbb{R}$ par*

$$f(x) := \max\{f_1(x), ..., f_m(x)\}$$

et on définit pour tout x l'ensemble des indices actifs $I(x) := \{i \mid f_i(x) = f(x)\}$. Alors,

$$\partial f(x) = conv \bigcup_{i \in I(x)} \partial f_i(x).$$

Corollaire 2.1.1 *Soient $f_1, ..., f_m : \mathbb{R}^n \to \mathbb{R}$ fonctions convexes et différentiables. On définit $f : \mathbb{R}^n \to \mathbb{R}$ par*

$$f(x) := \max\{f_1(x), ..., f_m(x)\}$$

et on définit pour tout x l'ensemble des indices actifs $I(x) := \{i \mid f_i(x) = f(x)\}$. Alors,

$$\partial f(x) = conv\,\{\nabla f_i(x) \mid i \in I(x)\}.$$

Nous terminons par une proposition qui sera utile dans une démonstration du chapitre 4.

Proposition 2.1.6 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe. La multifonction ∂f est monotone i.e.*
$\forall x_1, x_2 \in \mathbb{R}^n$, $\forall \xi_1 \in \partial f(x_1)$, $\forall \xi_2 \in \partial f(x_2)$ $\quad < \xi_2 - \xi_1, x_2 - x_1 > \geq 0.$

2.2 Méthode de plus forte pente

2.2.1 Généralités

On considère toujours une fonction $f : \mathbb{R}^n \to \mathbb{R}$ convexe et à valeurs finies. Le but de cette section est de présenter une méthode numérique pour trouver le minimum de f (s'il existe). Par méthode numérique, on sous-entend une procédure qui génère, à partir d'un point de départ x^0 , une suite

de points $\{x^k\}_{k\in\mathbb{N}}$ qui converge vers un minimum de f. La méthode est entièrement déterminée lorsque la façon dont on passe d'un point à l'autre est décrite, c'est à dire quand une itération est définie. La stratégie pour passer de x^k à x^{k+1} est la suivante :

1. définir une direction de recherche d^k,
2. poser $x^{k+1} = x^k + t^k d^k$ où $t^k > 0$ est une longueur de pas choisie intelligemment.

La direction d^k est habituellement choisie de façon à ce qu'il soit possible de diminuer la valeur de f le long de d^k en partant de x^k. Une telle direction est appelée direction de descente. En ce qui concerne le pas t^k, un choix théorique consiste à trouver celui qui minimise f le long de la direction d^k en résolvant le problème à une dimension :

$$\begin{array}{ll} \min & f(x^k + td^k) \\ \text{s.c.} & t > 0. \end{array}$$

Une telle procédure est appelée recherche linéaire exacte.
On peut également choisir un pas t^k permettant (si possible) une diminution "suffisante" de la valeur de f le long de d^k c'est à dire tel que

$$f(x^k + t^k d^k) < f(x^k) - \varepsilon \quad \text{où } \varepsilon > 0 \text{ fixé.}$$

Cette façon de procéder sera utilisée dans la section 2.4.2

2.2.2 Direction de descente

Puisque nous sommes à la recherche d'un minimum de f, il est logique d'exiger que la valeur de la fonction diminue au cours des itérations c'est-à-dire que $f(x^{k+1}) < f(x^k)$ (sauf si x^k est optimum). Puisque $x^{k+1} = x^k + t^k d^k$, on exigera que la direction d^k soit une direction de descente pour f en x^k ce qui signifie que pour des petits pas le long de d^k, la valeur de la fonction diminue. Plus précisément,

Définition 2.2.1 *Une direction $d \in \mathbb{R}^n$ est appelée direction de descente en x pour f si*

$$\exists\, \delta > 0 \text{ tel que } \forall t \in]0, \delta] \quad f(x + td) < f(x).$$

Puisque f est convexe, on peut démontrer que d est une direction de descente en x pour f si

$$\exists\, \delta > 0 \text{ tel que } f(x + \delta d) < f(x).$$

En effet, si $t \in]0, \delta]$ alors $\frac{t}{\delta} \in]0, 1]$ et par convexité de f, on a

$$\begin{aligned} f(x + td) &= f(\tfrac{t}{\delta}(x + \delta d) + (1 - \tfrac{t}{\delta})x) \\ &\leq \tfrac{t}{\delta} f(x + \delta d) + (1 - \tfrac{t}{\delta}) f(x) \\ &< \tfrac{t}{\delta} f(x) + f(x) - \tfrac{t}{\delta} f(x) \\ &= f(x). \end{aligned}$$

La proposition suivante caractérise le concept de direction de descente.

Proposition 2.2.1 *Les assertions suivantes sont équivalentes :*

1. *d est une direction de descente en x pour f ;*
2. *$f'(x, d) < 0$;*
3. *$< \xi, d > < 0 \quad \forall \xi \in \partial f(x)$.*

Quand f est différentiable en x, le sous-différentiel $\partial f(x)$ est réduit au singleton $\{\nabla f(x)\}$ et l'on retrouve le résultat bien connu :

$$d \text{ est une direction de descente en } x \text{ pour } f \iff \nabla f(x)^T d < 0.$$

De plus, dans ce cas, $d = -\nabla f(x)$ est une direction de descente en x pour f si $\nabla f(x) \neq 0$. Par contre, quand f est non différentiable en x, l'opposé d'un sous-gradient de f en x n'est pas nécessairement une direction de descente en x pour f.
La proposition suivante énonce une condition nécessaire et suffisante à l'existence d'une direction de descente.

Proposition 2.2.2 *Il existe une direction de descente en x pour f si et seulement si $0 \notin \partial f(x)$.*

Afin de généraliser la méthode de plus forte pente aux fonctions convexes non différentiables, nous considérons la "meilleure" direction de descente (localement).

Définition 2.2.2 *Soit $x \in \mathbb{R}^n$ tel que $0 \notin \partial f(x)$. La plus forte direction de descente en x pour f est définie par la solution du problème :*

$$\begin{array}{ll} \min_d & f'(x,d) \\ s.c. & \|d\| \leq 1, \end{array} \tag{2.1}$$

où $\|.\|$ est une norme sur $\mathbb{R}^n$.

Observons que le problème (2.1) admet toujours une solution car la dérivée directionnelle $f'(x,.)$ est une fonction continue. Puisque $f'(x,d) < 0$ pour un $d \in \mathbb{R}^n$ (car $0 \notin \partial f(x)$) et $f'(x,.)$ est positivement homogène c'est-à-dire $f'(x,td) = tf'(x,d)$ $\forall d \in \mathbb{R}^n$ et $t > 0$, on a que $\inf_{t>0} f'(x,td) = -\infty$. Pour obtenir une solution, il faut donc bien imposer une contrainte de norme sur d. Notons que la norme considérée dans cette section est la norme euclidienne classique.

Théorème 2.2.1 *Soit $x \in \mathbb{R}^n$ tel que $0 \notin \partial f(x)$. Alors,*

1. *le vecteur de norme minimale dans $\partial f(x)$ existe et est unique. Il coïncide avec la projection orthogonale de 0 sur $\partial f(x)$;*
2. *la plus forte direction de descente en x pour f est le vecteur $-m/\|m\|$ où m est le vecteur de norme minimale dans $\partial f(x)$.*

Si f est différentiable en x et $\nabla f(x) \neq 0$ alors $m = \nabla f(x)$ et on retrouve le résultat bien connu :

$d = -\nabla f(x)/\|\nabla f(x)\|$ est la plus forte direction de descente en x pour f.

2.2.3 Algorithme

On suppose dans cette section que pour tout x, il est possible de calculer la valeur de $f(x)$ et le sous-différentiel $\partial f(x)$ en entier. Quand f est différentiable, une itération de la méthode de plus forte pente consiste à se déplacer à partir de x^k le long de la plus forte direction de descente $d^k = -\nabla f(x^k)$ jusqu'à ce que le minimum soit atteint le long de cette direction. Quand f est convexe et non nécessairement différentiable, la méthode de plus forte pente devient :

Méthode de plus forte pente
0. Choisir un point de départ $x^0 \in \mathbb{R}^n$ et poser $k = 0$.
1. Calculer m i.e. le vecteur de norme minimal dans $\partial f(x^k)$.
2. Si $m = 0$ alors STOP $\Longrightarrow$ x_k minimum de f car $0 \in \partial f(x)$.
3. Poser $d^k = -m$ et calculer t^k solution du problème $\min_{t>0} f(x^k + td^k)$.
4. Poser $x^{k+1} = x^k + t^k d^k$.
5. Poser $k := k + 1$ et retourner à l'étape 1.

Exemple 2.2.1 *Soit* $f(x_1, x_2) = \max\{f_1(x_1, x_2), f_2(x_1, x_2), f_3(x_1, x_2)\}$ *où*

$$f_1(x_1, x_2) = x_1 + 2x_2 \quad f_2(x_1, x_2) = x_1 - 2x_2 \quad f_3(x_1, x_2) = -x_1.$$

Le minimum de f *est* $(0, 0)$. *Par le corollaire 2.1.1, le sous-différentiel en* (x_1, x_2) *est*

$$\partial f(x_1, x_2) = conv\,\{\nabla f_i(x_1, x_2) \mid i \in I(x_1, x_2)\}$$

où $I(x_1, x_2) = \{\, i \in \{1, 2, 3\} \mid f_i(x_1, x_2) = f(x_1, x_2)\}$. *Etant donné le point de départ* $(2, 2)^T$, *on a que* $I(x^0) = \{1\}$ *et* $\partial f(x^0) = \{\nabla f_1(x^0)\} = \{(1, 2)^T\}$. *Le vecteur de norme minimale est* $m = (1, 2)^T$ *et* $d^0 = -(1, 2)^T$. *La recherche linéaire exacte donne* $t^0 = 1$ *d'où* $x^1 = (2, 2)^T - (1, 2)^T = (1, 0)^T$. *A la seconde itération,* $I(x^1) = \{1, 2\}$ *et* $\partial f(x^1) = conv\,\{(1, 2)^T, (1, -2)^T\}$. *Le vecteur de norme minimale est* $m = (1, 0)^T$ *et* $d^1 = -(1, 0)^T$. *La recherche linéaire exacte donne de nouveau* $t^1 = 1$ *d'où* $x^2 = (0, 0)^T$. *Finalement,* $I(x^2) = \{1, 2, 3\}$ *et* $\partial f(x^2) = conv\,\{(1, 2)^T, (1, -2)^T, (-1, 0)^T\}$. *Le vecteur de norme minimale est* $m = (0, 0)^T$ *d'où* $x^2 = (0, 0)^T$ *est minimum de* f.

Dans cet exemple, la méthode converge vers le minimum. Néanmoins, il peut arriver que la méthode de plus forte pente converge vers un point non optimal. Plusieurs exemples sont présentés dans la littérature. En particulier, Lemaréchal a considéré une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x_1, x_2) = \max\{f_0(x_1, x_2), f_{\pm 1}(x_1, x_2), f_{\pm 2}(x_1, x_2)\}$$

où

$$f_0(x_1, x_2) = -100 \quad f_{\pm 1}(x_1, x_2) = 3x_1 \pm 2x_2 \quad f_{\pm 2}(x_1, x_2) = 2x_1 \pm 5x_2.$$

La valeur minimum de cette fonction est -100.
En partant de $x^0 = (9, -3)^T$, on a que $I(x^0) = \{-1, -2\}$ et $\partial f(x^0) =$ conv $\{(3, -2)^T, (2, -5)^T\}$. Le vecteur de norme minimale dans $\partial f(x^0)$ est

$m = (3, -2)^T$ d'où $d^0 = -(3, -2)^T$. La recherche linéaire exacte donne $t^0 = 2$ d'où $x^1 = (9, -3)^T - 2(3, -2)^T = (3, 1)^T$. A la seconde itération, $I(x^1) = \{1, 2\}$ et $\partial f(x^1) = \text{conv } \{(3, 2)^T, (2, 5)^T\}$. Le vecteur de norme minimale dans $\partial f(x^1)$ est $m = (3, 2)^T$ d'où $d^1 = -(3, 2)^T$. La recherche linéaire exacte donne $t^1 = 2/3$ d'où $x^2 = (1, -1/3)^T$. On obtient ensuite $x^3 = (1/3, 1/9)^T$ et le phénomène de zig-zag entre les deux rayons d'équations

$$\begin{aligned} x_1 \pm 3x_2 &= 0 \\ x_1 &\geq 0 \end{aligned}$$

devient évident. La suite $\{x^k\}$ converge lentement vers l'origine $x^* = (0, 0)^T$ qui n'est pas optimal. La non convergence de la suite $\{x^k\}$ est due à la discontinuité du sous-différentiel à l'origine. A chaque itération, le sous-différentiel est égal à conv $\{(3, -2)^T, (2, -5)^T\}$ ou à conv $\{(3, 2)^T, (2, 5)^T\}$ et ce, même si x^k est très proche de l'origine. Ensuite, le sous-différentiel devient brusquement à l'origine conv $\{(3, -2)^T, (2, -5)^T, (3, 2)^T, (2, 5)^T\}$. La méthode ne s'aperçoit pas de la discontinuité du sous-différentiel à la limite de la suite générée. Plus précisément, la propriété de continuité qui n'est pas vérifiée, en général, par le sous-différentiel est

$$\xi \in \partial f(x) \text{ et } x_i \to x \quad \Longrightarrow \quad \forall i \ \exists \, \xi_i \in \partial f(x_i) \text{ tel que } \xi_i \to \xi. \tag{2.2}$$

Afin de surpasser cette difficulté, Wolfe et Lemaréchal eurent l'idée d'agrandir le sous-différentiel $\partial f(x)$ en incluant des informations au sujet du voisinage de x (Proposition 2.3.1) afin que la propriété de continuité (2.2) soit vérifiée. Ce sera le sujet de la section suivante.

Finalement, remarquons que pour calculer le vecteur de norme minimale dans $\partial f(x)$, nous supposons que le sous-différentiel est connu dans son entièreté. Malheureusement, il est très souvent beaucoup trop coûteux (d'un point de vue numérique) de calculer tout le sous-différentiel. En témoigne l'exemple suivant :

Exemple 2.2.2 *Considèrons la fonction $\lambda_1 : S^m \to I\!R$. Remarquons que cette fonction est convexe et que son sous-différentiel en $M \in S^m$ est donné par*

$$\partial\lambda_1(M) = \textit{conv } \{qq^T \mid q^Tq = 1, \ Mq = \lambda_1(M)q\}.$$

Par conséquent, si on désire calculer le sous-différentiel de λ_1 en M, il est nécessaire de calculer tous les vecteurs propres associés à $\lambda_1(M)$ ce qui est trop coûteux. Heureusement, calculer un sous-gradient l'est moins car cela revient seulement à déterminer un vecteur propre normalisé associé à $\lambda_1(M)$.

2.3 Sous-gradient et sous-différentiel approximés

Le concept de sous-gradient peut lui aussi être généralisé par la pente d'une fonction affine minorant f mais non nécessairement égale à f en un point. Cette généralisation porte le nom de sous-gradient approximé ou ε-sous-gradient.

Définition 2.3.1 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe, $x \in \mathbb{R}^n$ et $\varepsilon \geq 0$. Alors, $\xi \in \mathbb{R}^n$ est un ε-sous-gradient de f en x si*

$$f(y) \geq f(x) + < \xi, y - x > -\varepsilon \quad \forall y \in \mathbb{R}^n.$$

L'ensemble des ε-sous-gradients de f en x est appelé ε-sous-différentiel de f en x et est noté $\partial_\varepsilon f(x)$.

On a immédiatement que

$$\partial_\varepsilon f(x) \subseteq \partial_{\varepsilon'} f(x) \text{ pour } \varepsilon \leq \varepsilon'$$

et

$$\partial_0 f(x) = \partial f(x) = \cap_{\varepsilon > 0} \partial_\varepsilon f(x).$$

Les fonctions affines qui correspondent aux sous-gradients en x doivent toucher la fonction en x alors que les fonctions affines qui correspondent aux ε-sous-gradients en x se trouvent en dessous de la fonction en x à une distance ε. Intuitivement, on remarque que $\partial_\varepsilon f(x)$ capture davantage le comportement local de f dans un voisinage du point x que ne le fait $\partial f(x)$ et ce, au prix d'une certaine perte de précision. En fait, $\partial_\varepsilon f(x)$ contient tous les sous-différentiels d'une boule centrée en x.

Proposition 2.3.1 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe. Pour tout x et tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que*

$$\partial_\varepsilon f(x) \supset \bigcup_{y \in B(x,\delta)} \partial f(y).$$

Tout comme le sous-différentiel, le sous-différentiel approximé est un ensemble non vide, convexe et compact. Il admet des propriétés similaires à

celles du sous-différentiel.
La propriété suivante n'est valable que pour l'ε-sous-différentiel où $\varepsilon > 0$.

Proposition 2.3.2 *Soient $\{x_k\}_k \subset \mathbb{R}^n$ une suite convergeant vers $x \in \mathbb{R}^n$ et $\xi \in \partial f_\varepsilon(x)$. Alors, il existe une suite $\{\xi_k\}_k$ convergeant vers ξ telle que $\xi_k \in \partial f_\varepsilon(x_k)$ pour tout k.*

Rappelons que cette propriété n'est pas toujours satisfaite par le sous-différentiel. C'est l'une des raisons pour laquelle la méthode de plus forte pente peut ne pas converger.
On peut définir la notion d'optimalité approximée de deux façons :

Définition 2.3.2 *Un point $x^* \in \mathbb{R}^n$ est dit ε-optimal si*

$$f(x) \geq f(x^*) - \varepsilon \quad \forall x \in \mathbb{R}^n.$$

La proposition suivante est analogue à la Proposition 2.1.3.

Proposition 2.3.3 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe, $x^* \in \mathbb{R}^n$ et $\varepsilon \geq 0$. Alors, x^* est un point ε-optimal si et seulement si $0 \in \partial_\varepsilon f(x^*)$.*

Définition 2.3.3 *Un point $x^* \in \mathbb{R}^n$ est dit ε-stationnaire si*

$$\exists\, \xi \in \partial_\varepsilon f(x^*) \text{ tel que } \|\xi\| \leq \varepsilon.$$

En réalité, un point ε-stationnaire x peut être arbitrairement éloigné de l'optimum x^*. En d'autres termes, $|f(x) - f(x^*)|$ et $\|x - x^*\|$ peuvent être arbitrairement grand. Cependant, le concept de point ε-stationnaire est utile car il fournit un critère d'arrêt pour les méthodes faisceaux.
La proposition suivante est utile dans les preuves de convergence d'algorithmes utilisant des sous-gradients approximés.

Proposition 2.3.4 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ fonction convexe. Soient $\{x^i\}$, $\{\xi^i\}$ des suites dans $\mathbb{R}^n$ et $\{\varepsilon^i\}$ une suite de réels positifs. Si $x^i \to x$, $\xi^i \in \partial_{\varepsilon^i} f(x^i)$, $\xi^i \to 0$ et $\varepsilon_i \to 0$ alors, x minimise f.*

Les règles de calcul pour le sous-différentiel approximé sont semblables à celles du sous-différentiel.

Proposition 2.3.5 *Soient $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ fonctions convexes et $\varepsilon \geq 0$. On définit $f : \mathbb{R}^n \to \mathbb{R}$ par*

$$f(x) := f_1(x) + f_2(x).$$

Alors,

$$\partial_\varepsilon f(x) = \bigcup \{\partial_{\varepsilon_1} f_1(x) + \partial_{\varepsilon_2} f_2(x) \mid \varepsilon_1, \varepsilon_2 \geq 0,\ \varepsilon_1 + \varepsilon_2 \leq \varepsilon\}.$$

Proposition 2.3.6 *Soient $f_1, ..., f_m : \mathbb{R}^n \to \mathbb{R}$ fonctions convexes et $\varepsilon \geq 0$. On définit $f : \mathbb{R}^n \to \mathbb{R}$ par*

$$f(x) := \max\{f_1(x), ..., f_m(x)\}.$$

Alors,

$$\partial_\varepsilon f(x) = \bigcup \left\{ \sum_{i=1}^m \alpha_i \xi_i \mid \alpha \in \Delta_m,\ \xi_i \in \partial_{\varepsilon_i/\alpha_i} f_i(x) \text{ si } \alpha_i > 0,\ \sum_{i=1}^m (\varepsilon_i + \alpha_i e_i) \leq \varepsilon \right\} \tag{2.3}$$

où Δ_m est le simplexe unité $\{\alpha \in \mathbb{R}^m \mid \alpha_i \geq 0,\ \sum_{i=1}^m \alpha_i = 1\}$ et $e_i := f(x) - f_i(x)$.

Les sous-gradients approximés en un point peuvent être obtenus de multiples façons. Une façon de faire consiste à "transporter" un sous-gradient exact ou approximé d'un autre point.

Proposition 2.3.7 (Formule de transfert) *Soient $f : \mathbb{R}^n \to \mathbb{R}$ fonction convexe, $\varepsilon \geq 0$ et $x, y \in \mathbb{R}^n$. Alors,*

$$\xi \in \partial_\varepsilon f(y) \iff \xi \in \partial_{\varepsilon+\alpha(x,y)} f(x) \tag{2.4}$$

où

$$\alpha(x,y) := f(x) - f(y) - < \xi, x - y > .$$

La proposition décrit le transfert d'un sous-gradient approximé en y en un sous-gradient approximé en x c'est pourquoi (2.4) est appelée formule de transfert. La quantité $\alpha(x,y)$ est appelée erreur de linéarisation car elle exprime l'erreur entre la valeur de la fonction en x et la valeur de la linéarisation de la fonction (définie par un sous-gradient ξ en y) en x.

2.4 Méthode d' ε plus forte pente

2.4.1 Direction d' ε-descente

Pour présenter la méthode d' ε plus forte pente, nous avons besoin de la définition suivante. Une direction $d \in \mathbb{R}^n$ est une direction d'ε-descente en x pour f où $\varepsilon > 0$ s'il est possible de diminuer la valeur de la fonction d'une valeur ε en partant de x le long de d. Plus précisément,

Définition 2.4.1 *Une direction $d \in \mathbb{R}^n$ est appelée direction d' ε-descente en x pour f si*

$$\exists\, t > 0 \quad \textit{tel que} \quad f(x + td) < f(x) - \varepsilon.$$

La proposition suivante caractérise le concept de direction d' ε-descente.

Proposition 2.4.1 *Les assertions suivantes sont équivalentes :*

1. *d est une direction d' ε-descente en x pour f ;*
2. *$< \xi, d > < 0 \quad \forall \xi \in \partial f_\varepsilon(x)$.*

La proposition suivante énonce une condition nécessaire et suffisante d'existence d'une direction d' ε-descente.

Proposition 2.4.2 *Les propriétés suivantes sont vérifiées :*

1. $0 \notin \partial f_\varepsilon(x) \iff \exists\, d \in \mathbb{R}^n$ *tel que* d *est une direction d'* ε*-descente en* x *pour* f.
2. $0 \notin \partial f_\varepsilon(x) \implies d = -Proj\ 0/\partial_\varepsilon f(x)$ *est une direction d'* ε*-descente en* x *pour* f.

2.4.2 Algorithme

On considère toujours une fonction $f : \mathbb{R}^n \to \mathbb{R}$ convexe, à valeurs finies et non nécessairement différentiable. On suppose dans cette section que pour tout x et pour $\varepsilon > 0$ fixé, il est possible de calculer la valeur de $f(x)$ et l'ε-sous-différentiel $\partial_\varepsilon f(x)$ en entier.

Méthode d' ε plus forte pente

0. Choisir un point de départ $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$ et poser $k = 0$.
1. Calculer le vecteur m de norme minimale dans $\partial f_\varepsilon(x^k)$.
2. Si $m = 0$ alors STOP $\implies$ x^k est un ε-minimum de f car $0 \in \partial_\varepsilon f(x^k)$.
3. Poser $d^k = -m$ et calculer t^k tel que $f(x^k + t^k d^k) < f(x^k) - \varepsilon$.
4. Poser $x^{k+1} = x^k + t^k d^k$.
5. Poser $k := k + 1$ et retourner à l'étape 1.

Contrairement à la méthode de plus forte pente, la méthode d' ε plus forte pente converge. En témoigne le théorème suivant.

Théorème 2.4.1 *Soit* $\{x^k\}_k$ *la suite générée par la méthode d'* ε *plus forte pente. Alors,* $f(x^k) \to -\infty$ *ou le processus d'optimisation s'arrête au Pas 2 après un nombre fini d'itérations donnant lieu a un* ε*-minimum de* f.

Les inconvénients de la méthode d' ε plus forte pente sont premièrement, son manque de flexibilité. En effet, l'objectif est de diminuer à chaque itération la valeur de f d'au moins une quantité $\varepsilon > 0$ fixée à priori et ce, quel que soit

le comportement de f. Il serait plus intéressant de pouvoir changer la valeur de ε à chaque itération. Le second inconvénient réside dans les exigences de la méthode. En effet, comme pour la méthode de plus forte pente, le sous-différentiel approximé de chaque itéré est supposé connu.

2.5 Méthode du sous-gradient

Comme nous l'avons vu dans les sections précédentes, la méthode de plus forte pente et la méthode d' ε plus forte pente nécessitent la connaissance respectivement du sous-différentiel et de l'ε-sous-différentiel à chaque itération. Malheureusement, de telles connaissances ne sont pas souvent disponibles en pratique. Cependant, dans la plupart des exemples pratiques, on peut supposer l'existence d'une boîte noire qui à chaque point $x \in \mathbb{R}^n$ renvoie la valeur $f(x)$ et un sous-gradient $\xi \in \partial f(x)$. Une telle boîte noire est souvent appelée oracle. Par exemple, un oracle pour une fonction max

$$f(x) = \max\{f_1(x), f_2(x), ..., f_m(x)\}$$

où les f_i sont différentiables consiste (par le corollaire 2.1.1) à trouver un indice $i \in \{1, ..., m\}$ tel que $f(x) = f_i(x)$ et à renvoyer $f_i(x)$ ainsi que $\nabla f_i(x)$ (un sous-gradient de f en x).

Puisqu'on considère une fonction $f : \mathbb{R}^n \to \mathbb{R}$ convexe, à valeurs finies et par conséquent différentiable presque partout, le sous-différentiel est réduit au gradient presque partout. On peut donc simplement prétendre que la fonction est différentiable et le sous-gradient donné par l'oracle peut jouer le rôle du gradient dans la méthode de plus forte pente classique (cas différentiable). Plus précisément, l'itération pend la forme

$$x^{k+1} = x^k - t^k \xi^k$$

où $\xi^k \in \partial f(x^k)$ est donné par l'oracle et $t^k > 0$ est une longueur de pas. Cette méthode, appelée méthode du sous-gradient, a quelques inconvénients. D'une part, comme nous l'avons vu dans la section 2.2.2, l'opposé d'un sous-gradient n'est en général pas une direction de descente c'est pourquoi sélectionner la longueur de pas est très difficile. D'autre part, il n'existe pas de critère d'arrêt pratique car l'oracle peut retourner n'importe quel sous-gradient (même au minimum, un sous-gradient différent de zéro peut-être obtenu).

Chapitre 3

Méthodes Faisceaux

Ce chapitre introduit une famille d'algorithmes, appelés méthodes faisceaux, adaptés à la minimisation de fonctions convexes non différentiables et il présente les points de vue primal et dual de ces méthodes. Tout comme la méthode du sous-gradient, les méthodes faisceaux sont construites à partir de l'hypothèse suivante :

Hypothèse 3.0.1 *Une sous-routine qui pour tout x calcule $f(x)$ et un sous-gradient $\xi \in \partial f(x)$ est disponible.*

Les preuves de convergence sont présentées au chapitre suivant puisque ce dernier décrit une classe plus générale d'algorithmes à laquelle les méthodes faisceaux appartiennent.

3.1 Point de vue dual des méthodes faisceaux

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe. Nous avons vu dans la section 2.4.1 qu'une direction d' ε-descente en x pour f se calculait en cherchant le vecteur de norme minimale dans $\partial_\varepsilon f(x)$ c'est à dire en résolvant le problème

$$\begin{array}{ll} \min & \|\xi\| \\ \text{s.c.} & \xi \in \partial_\varepsilon f(x). \end{array}$$

Puisque l'entièreté de l'ε-sous-différentiel n'est en général pas disponible en x, nous allons construire une approximation de $\partial_\varepsilon f(x)$ à l'aide de l'oracle disponible par l'Hypothèse 3.0.1. L'approximation sera construite à l'aide de sous-gradients calculés dans un voisinage de x. Plus précisément, supposons

que x^k est le point d'itération actuel et que pendant le processus d'optimisation, nous avons obtenu les points $y^j \in \mathbb{R}^n$, $j = 1, ..., k$ et les informations correspondantes $f(y^j)$ et $\xi^j \in \partial f(y^j)$ grâce à l'oracle. Alors, par la formule de transfert (Proposition 2.3.7), nous avons que

$$\xi^j \in \partial_{\alpha_j^k} f(x^k) \tag{3.1}$$

où $\alpha_j^k = \alpha(x^k, y^j) = f(x^k) - f(y^j) - < \xi^j, x^k - y^j >$ est l'erreur de linéarisation. Soit $J_k \subseteq \{1, ..., k\}$. L'ensemble $\{(\xi^j, \alpha_j^k)\}_{j \in J_k}$ est appelé faisceau. Le faisceau représente une collection de sous-gradients approximés disponible au point d'itération actuel x^k. La proposition suivante nous montre que le faisceau nous permet de construire une approximation intérieure de $\partial_\varepsilon f(x^k)$.

Proposition 3.1.1 *L'ensemble*

$$G(x^k, \varepsilon) = \left\{ \sum_{j \in J_k} \lambda_j \xi^j \mid \lambda_j \geq 0,\ j \in J_k,\ \sum_{j \in J_k} \lambda_j = 1,\ \sum_{j \in J_k} \lambda_j \alpha_j^k \leq \varepsilon \right\}$$

est un sous-ensemble convexe de $\partial_\varepsilon f(x^k)$.

Preuve :

1. $G(x^k, \varepsilon)$ est convexe car $\forall \sum_{j \in J_k} \lambda_j \xi^j,\ \sum_{j \in J_k} \mu_j \xi^j \in G(x^k, \varepsilon)$ et $\forall \beta \in]0, 1[$, on a que

$$\beta \sum_{j \in J_k} \lambda_j \xi^j + (1 - \beta) \sum_{j \in J_k} \mu_j \xi^j = \sum_{j \in J_k} (\beta \lambda_j + (1 - \beta) \mu_j) \xi^j$$

tel que

$\beta \lambda_j + (1 - \beta) \mu_j \geq 0,$

$$\begin{aligned} \textstyle\sum_{j \in J_k} (\beta \lambda_j + (1 - \beta) \mu_j) &= \beta \textstyle\sum_{j \in J_k} \lambda_j + (1 - \beta) \sum_{j \in J_k} \mu_j \\ &= \beta + 1 - \beta \\ &= 1, \\ \textstyle\sum_{j \in J_k} (\beta \lambda_j + (1 - \beta) \mu_j) \alpha_j^k &= \beta \textstyle\sum_{j \in J_k} \lambda_j \alpha_j^k + (1 - \beta) \sum_{j \in J_k} \mu_j \alpha_j^k \\ &\leq \beta \varepsilon + (1 - \beta) \varepsilon \\ &= \varepsilon. \end{aligned}$$

2. Par (3.1), $\xi^j \in \partial_{\alpha_j^k} f(x^k)$ $\forall j \in J_k$ c'est à dire

$$\forall j \in J_k,\ \forall y \in \mathbb{R}^n \quad f(y) \geq f(x^k) + < \xi^j, y - x^k > -\alpha_j^k.$$

Donc, pour $\lambda_j \geq 0$, $j \in J_k$ tels que $\sum_{j \in J_k} \lambda_j = 1$ et $\sum_{j \in J_k} \lambda_j \alpha_j^k \leq \varepsilon$, nous avons que

$$\forall j \in J_k,\ \forall y \in \mathbb{R}^n \quad \lambda_j f(y) \geq \lambda_j f(x^k) + < \lambda_j \xi^j, y - x^k > -\lambda_j \alpha_j^k.$$

En sommant les $|J_k|$ équations, on obtient $\forall y \in \mathbb{R}^n$

$$\sum_{j \in J_k} \lambda_j f(y) \geq \sum_{j \in J_k} \lambda_j f(x^k) + < \sum_{j \in J_k} \lambda_j \xi^j, y - x^k > - \sum_{j \in J_k} \lambda_j \alpha_j^k$$

ou encore

$$\forall y \in \mathbb{R}^n \quad f(y) \geq f(x^k) + < \sum_{j \in J_k} \lambda_j \xi^j, y - x^k > - \sum_{j \in J_k} \lambda_j \alpha_j^k.$$

On obtient alors que $\sum_{j \in J_k} \lambda_j \xi^j \in \partial_{\sum_{j \in J_k} \lambda_j \alpha_j^k} f(x^k)$ et puisque $\sum_{j \in J_k} \lambda_j \alpha_j^k \leq \varepsilon$, on obtient finalement que

$$\sum_{j \in J_k} \lambda_j \xi^j \in \partial_{\sum_{j \in J_k} \lambda_j \alpha_j^k} f(x^k) \subseteq \partial_\varepsilon f(x^k). \tag{3.2}$$

On peut alors conclure que $G(x^k, \varepsilon) \subseteq \partial_\varepsilon f(x^k)$.

□

Grâce à cette proposition, on peut construire une approximation de la plus forte direction d' ε-descente en calculant le vecteur de norme minimale dans $G(x^k, \varepsilon)$.

Approximation de la plus forte direction d' ε descente

1. Résoudre le problème quadratique convexe

$$\begin{array}{ll} \min & \frac{1}{2} \| \sum_{j \in J_k} \lambda_j \xi^j \|^2 \\ \text{s.c.} & \lambda_j \geq 0 \quad \forall j \in J_k \\ & \sum_{j \in J_k} \lambda_j = 1 \\ & \sum_{j \in J_k} \lambda_j \alpha_j^k \leq \varepsilon \end{array} \tag{3.3}$$

pour obtenir la solution λ_j^*, $j \in J_k$.

2. Poser $d_k = -\sum_{j\in J_k} \lambda_j^* \xi^j$.

Remarquons que le problème (3.3) admet une solution si son ensemble admissible est non vide c'est à dire si $\min_{j\in J_k} \alpha_j^k \leq \varepsilon$. Par ailleurs, puisque $G(x^k, \varepsilon)$ n'est qu'une approximation de $\partial_\varepsilon f(x^k)$, on ne peut garantir que d^k est une direction de descente en x^k pour f. On ne peut donc pas utiliser une recherche linéaire classique pour trouver x^{k+1} c'est pourquoi on utilise une recherche linéaire à deux sorties. La première sortie (appelée pas sérieux) correspond au cas où $G(x^k, \varepsilon)$ est une bonne approximation de $\partial_\varepsilon f(x^k)$ et la seconde sortie (appelée pas nul) correspond au cas où $G(x^k, \varepsilon)$ est une mauvaise approximation de $\partial_\varepsilon f(x^k)$. Plus précisément, un pas $t > 0$ est appelé pas sérieux si la diminution de f : $f(x^k) - f(x^k + td^k)$ est suffisamment grande et si la longueur de pas t n'est pas trop petite. Dans ce cas, x^k est mis à jour : $x^{k+1} = x^k + td^k$. Par contre, si en diminuant la valeur de t, ces deux conditions ne sont jamais vérifiées, alors le pas est nul. L'itéré x^k n'est pas mis à jour : $x^{k+1} = x^k$ et le sous-gradient obtenu en $x^k + td^k$ pour t petit est ajouté au faisceau afin que l'approximation de $\partial_\varepsilon f(x^k)$ soit de meilleure qualité à l'itération suivante. Cette stratégie est assez compliquée et difficile à implémenter. De plus, en pratique, aucun élément nous permet de choisir à raison dans (3.3) une valeur particulière de ε. En effet, la valeur de ε doit refléter la possible diminution de f en x^k qui est inconnue à priori. Par conséquent, nous résoudrons un autre problème quadratique convexe à la place de (3.3) afin de calculer d^k :

Approximation de la plus forte direction d' ε descente

1. Résoudre le problème quadratique convexe

$$\begin{array}{ll} \min & \frac{1}{2}\|\sum_{j\in J_k} \lambda_j \xi^j\|^2 + u\sum_{j\in J_k} \lambda_j \alpha_j^k \\ s.c. & \lambda_j \geq 0 \quad \forall j \in J_k \\ & \sum_{j\in J_k} \lambda_j = 1 \end{array} \tag{3.4}$$

   pour $u \geq 0$ afin d'obtenir la solution λ_j^*, $j \in J_k$.

2. Poser $d_k = -\sum_{j\in J_k} \lambda_j^* \xi^j$.

Les deux problèmes quadratiques (3.3) et (3.4) sont équivalents dans le sens de la proposition suivante dont la preuve est disponible dans [5].

Proposition 3.1.2 *Les propriétés suivantes sont vérifiées :*

1. *Si* $\{\lambda_j^k\}_{j\in J_k}$ *est solution optimale de (3.3) et si* $u_k \geq 0$ *correspond au multiplicateur de Lagrange associé à la contrainte*

$$\sum_{j\in J_k} \lambda_j \alpha_j^k \leq \varepsilon,$$

alors $\{\lambda_j^k\}_{j\in J_k}$ *est aussi solution optimale de (3.4) pour* $u = u_k$.

2. *Si* $\{\lambda_j^k\}_{j\in J_k}$ *est solution optimale de (3.4), alors* $\{\lambda_j^k\}_{j\in J_k}$ *est aussi solution optimale de (3.3) où*

$$\varepsilon = \sum_{j\in J_k} \lambda_j^k \alpha_j^k.$$

Remarquons qu'à la place de choisir ε dans (3.3), le problème revient maintenant à choisir la valeur de u dans (3.4). En pratique, les méthodes faisceaux (sous la forme duale) résolvent le problème (3.4) pour trouver la direction de recherche d^k.
Notons que le vecteur $\bar{\xi}^k = \sum_{j\in J_k} \lambda_j^k \xi^j$ où $\{\lambda_j^k\}_{j\in J_k}$ est solution du problème (3.4) jouera un rôle pratique central dans les méthodes faisceaux. Nous avons par (3.2) que $\bar{\xi}^k \in \partial_{\bar{\alpha}^k} f(x^k)$ où $\bar{\alpha}^k = \sum_{j\in J_k} \lambda_j^k \alpha_j^k$. Ce sous-gradient approximé est appelé sous-gradient agrégé. Il nous permettra de limiter la taille du faisceau.
Finalement, mentionnons que les méthodes faisceaux actuelles réduisent la recherche linéaire en une décision de prendre ou non un pas de longueur fixée dans la direction de d^k. Si la diminution $f(x^k) - f(x^k + d^k)$ est suffisante alors le pas est sérieux et $x^{k+1} = x^k + d^k$. Sinon, le pas est nul et $x^{k+1} = x^k$. Il est plus simple de décrire ceci à partir de l'approche primale des méthodes faisceaux.

3.2 Point de vue primal des méthodes faisceaux

Dans cette section, nous n'utilisons plus les sous-gradients pour approximer $\partial_\varepsilon f(x)$ mais pour définir des fonctions affines minorant f. Tout comme dans la section précédente, supposons que x^k est le point d'itération actuel et que pendant le processus d'optimisation, nous avons obtenu les points $y^j \in \mathbb{R}^n$, $j = 1, ..., k$ et les informations correspondantes $f(y^j)$ et $\xi^j \in \partial f(y^j)$

grâce à l'oracle. Par définition d'un sous-gradient en y^j pour f, nous avons que

$$f(x) \geq f(y^j) + < \xi^j, x - y^j > \quad \forall x \in \mathbb{R}^n.$$

Par conséquent, la fonction convexe définie pour $x \in \mathbb{R}^n$ par

$$\varphi^k(x) = \max \{f(y^j) + < \xi^j, x - y^j > \mid j \in J_k\}$$

minore f où $J_k \subseteq \{1, ...k\}$. Cette fonction est appelée modèle du plan sécant. En utilisant l'expression de l'erreur de linéarisation

$$\alpha_j^k = f(x^k) - f(y^j) - < \xi^j, x^k - y^j > \quad j \in J_k,$$

le modèle de plan sécant peut être réécrit en terme de x^k

$$\begin{aligned} \varphi^k(x) &= \max \{f(x^k) - < \xi^j, x^k - y^j > -\alpha_j^k + < \xi^j, x - y^j > \mid j \in J_k\} \\ &= \max \{f(x^k) + < \xi^j, x - x^k > -\alpha_j^k \mid j \in J_k\}, \end{aligned}$$

ou encore

$$\varphi^k(x) = f(x^k) + \max \{< \xi^j, x - x^k > -\alpha_j^k \mid j \in J_k\}. \tag{3.5}$$

En utilisant ce modèle, une direction de recherche en x^k pour f est obtenue en résolvant le problème

$$\begin{aligned} d^k &= \operatorname{argmin}_d \{\varphi^k(x^k + d) - f(x^k)\} \\ &= \operatorname{argmin}_d \max \{< \xi^j, d > -\alpha_j^k \mid j \in J_k\}. \end{aligned} \tag{3.6}$$

En ajoutant une variable, le problème (3.6) peut être reformulé en un problème de programmation linéaire ordinaire

$$\begin{aligned} &\min_{v,d} \quad v \\ &\text{s.c.} \quad\quad < \xi^j, d > -\alpha_j^k \leq v \quad j \in J_k. \end{aligned} \tag{3.7}$$

L'inconvénient du modèle du plan sécant est que la solution du problème (3.6) peut ne pas être bornée. Pour éviter cela, une stratégie consiste à ajouter une contrainte de norme sur d dans (3.7). Le problème devient alors

$$\begin{aligned} &\min_{v,d} \quad v \\ &\text{s.c.} \quad\quad < \xi^j, d > -\alpha_j^k \leq v \quad j \in J_k \\ &\qquad\quad \tfrac{1}{2}\|d\|^2 \leq \rho, \end{aligned} \tag{3.8}$$

où $\rho > 0$ est le rayon de la région de confiance. Une autre stratégie consiste à ajouter un terme quadratique à la fonction objectif du problème (3.7).

Ceci permet d'éviter l'apparition de la contrainte quadratique. Le problème devient alors

$$\begin{array}{ll} \min_{v,d} & v + \frac{1}{2}u\|d\|^2 \\ \text{s.c.} & < \xi^j, d > -\alpha_j^k \leq v \quad j \in J_k, \end{array} \tag{3.9}$$

où $u > 0$. On peut démontrer que les problèmes (3.8) et (3.9) sont équivalents de la manière suivante :

- Si (v^*, d^*) est solution de (3.8) et si u^* est le multiplicateur de Lagrange associé à la contrainte quadratique alors (v^*, d^*) est solution de (3.9) où $u = u^*$.
- Si (v^*, d^*) est solution de (3.9) alors (v^*, d^*) est solution de (3.8) où $\rho = \frac{1}{2}\|d^*\|^2$.

Puisqu'il est difficile de choisir la valeur de ρ dans le problème (3.8), les méthodes faisceaux (approche primale) utilise le problème (3.9) pour calculer la direction de recherche d^k. L'intérêt de l'utilisation du problème (3.9) réside également dans l'expression de son dual Lagrangien. En effet, comme le montre la proposition suivante, le problème dual de (3.9) n'est autre que le problème (3.4) qui consistait à calculer une approximation de la plus forte direction d' ε descente.

Proposition 3.2.1 *Le dual Lagrangien du problème (3.9) est*

$$\begin{array}{ll} \min & \frac{1}{2}\|\sum_{j\in J_k} \lambda_j \xi^j\|^2 + u\sum_{j\in J_k} \lambda_j \alpha_j^k \\ s.c. & \lambda_j \geq 0 \quad \forall j \in J_k \\ & \sum_{j\in J_k} \lambda_j = 1. \end{array}$$

Si $\{\lambda_j^k\}_{j\in J_k}$ est solution du dual lagrangien alors la solution (v^k, d^k) de (3.9) est donnée par

$$v^k = -\frac{1}{u}\|\bar{\xi}^k\|^2 - \bar{\alpha}^k \text{ et } d^k = -\frac{1}{u}\sum_{j\in J_k} \lambda_j^k \xi^j,$$

où $\bar{\xi}^k = \sum_{j\in J_k} \lambda_j^k \xi^j$ et $\bar{\alpha}^k = \sum_{j\in J_k} \lambda_j^k \alpha_j^k$.

Preuve :

1. Le Lagrangien associé au problème (3.9) est

$$L(v,d,\lambda) = v + \frac{1}{2}u\|d\|^2 + \sum_{j\in J_k} \lambda_j(< \xi^j, d > -\alpha_j^k - v).$$

La fonction duale est $\theta(\lambda) = \min_{v,d} L(v,d,\lambda)$. Pour trouver le minimum de $L(v,d,\lambda)$, on résout le système

$$\begin{array}{ll} \nabla_v L(v,d,\lambda) & = 1 - \sum_{j\in J_k} \lambda_j \quad = 0, \\ \nabla_d L(v,d,\lambda) & = ud + \sum_{j\in J_k} \lambda_j \xi^j \quad = 0. \end{array}$$

$v = 0$ et $d = -\frac{1}{u}\sum_{j\in J_k} \lambda_j \xi^j$ est solution du système. La fonction duale devient alors

$$\begin{aligned} \theta(\lambda) &= \textstyle\sum_{j\in J_k} \lambda_j(< \xi^j, -\frac{1}{u}\sum_{j\in J_k} \lambda_j \xi^j > -\alpha_j^k) \\ &\quad + \tfrac{1}{2}u\| - \tfrac{1}{u}\textstyle\sum_{j\in J_k} \lambda_j \xi^j\|^2 \\ &= \tfrac{1}{2u}\|\textstyle\sum_{j\in J_k} \lambda_j \xi^j\|^2 - \tfrac{1}{u} < \textstyle\sum_{j\in J_k} \lambda_j \xi^j, \sum_{j\in J_k} \lambda_j \xi^j > \\ &\quad - \textstyle\sum_{j\in J_k} \lambda_j \alpha_j^k \\ &= -\tfrac{1}{2u}\|\textstyle\sum_{j\in J_k} \lambda_j \xi^j\|^2 - \sum_{j\in J_k} \lambda_j \alpha_j^k. \end{aligned}$$

Par conséquent, le problème dual est

$$\begin{array}{ll} \max & -\frac{1}{2u}\|\sum_{j\in J_k} \lambda_j \xi^j\|^2 - \sum_{j\in J_k} \lambda_j \alpha_j^k \\ \text{s.c.} & \sum_{j\in J_k} \lambda_j = 1 \\ & \lambda_j \geq 0 \quad \forall j \in J_k, \end{array}$$

ou encore

$$\begin{array}{ll} -\min & \frac{1}{2u}\|\sum_{j\in J_k} \lambda_j \xi^j\|^2 + \sum_{j\in J_k} \lambda_j \alpha_j^k \\ \text{s.c.} & \sum_{j\in J_k} \lambda_j = 1 \\ & \lambda_j \geq 0 \quad \forall j \in J_k. \end{array} \tag{3.10}$$

Après multiplication par $u > 0$ de la fonction objectif de (3.10), on obtient l'expression désirée du problème dual.

2. En utilisant la condition de complémentarité, nous avons

$$\sum_{j\in J_k} \lambda_j^k (< \xi^j, d^k > -\alpha_j^k - v^k) = 0.$$

Or,

$$\begin{aligned} v^k &= \underbrace{\sum_{j\in J_k} \lambda_j^k v^k}_{=1} = \textstyle\sum_{j\in J_k} \lambda_j^k < \xi^j, d^k > - \sum_{j\in J_k} \lambda_j^k \alpha_j^k \\ &= < \bar{\xi}^k, -\tfrac{1}{u} \textstyle\sum_{j\in J_k} \lambda_j^k \xi^j > - \bar{\alpha}^k \\ &= -\tfrac{1}{u}\|\bar{\xi}^k\|^2 - \bar{\alpha}^k. \end{aligned}$$

□

On déduit immédiatement de (3.9) que la solution en v est $v^k = \varphi^k(x^k + d^k) - f(x^k)$ c'est à dire la diminution attendue de f lorsque l'on se déplace de x^k à $x^k + d^k$. Remarquons que v^k est nécessairement négatif car $\bar{\alpha}^k$ est positif par définition.

- Si $v^k = 0$ alors x^k minimise f car $\bar{\xi}^k \in \partial_{\bar{\alpha}^k} f(x^k)$ et

$$\underbrace{-\frac{1}{u}\|\bar{\xi}^k\|^2}_{\leq 0} = \underbrace{\bar{\alpha}^k}_{\geq 0} \quad \Longrightarrow \quad \bar{\xi}^k = \bar{\alpha}^k = 0,$$

  d'où $0 \in \partial f(x^k)$.

- Si $v^k < 0$ alors on s'attend à ce que d^k soit une direction de descente. Cependant, Il peut arriver que cela ne soit pas le cas puisque l'on travaille avec φ^k, une approximation de f. La stratégie des méthodes faisceaux est alors la suivante :

  - On pose $y^{k+1} = x^k + d^k$ le nouveau point de test.

  - Si $f(y^{k+1}) < f(x^k) + mv^k$ pour un $m \in (0,1)$ alors la diminution de f est jugée suffisante et un pas sérieux est pris en posant $x^{k+1} = y^{k+1}$.

  - Sinon, la diminution n'est pas jugée suffisante et un pas nul est pris en posant $x^{k+1} = x^k$. Dans ce cas, le nouveau modèle de plan sécant φ^{k+1} doit inclure le sous-gradient en y^{k+1} i.e. $k+1 \in J_{k+1}$ afin d'obtenir une meilleure direction de recherche à l'itération suivante.

Remarquons que la procédure dans son entièreté peut être interprétée en terme de deux boucles imbriquées : une boucle intérieure définie par des pas nuls consécutifs qui est à la recherche d'une direction de descente adaptée et une boucle extérieure définie par des pas sérieux qui produisent la diminution de la fonction objectif.
Pour clarifier les idées décrites précédemment, nous résumons la méthode faisceau par l'algorithme suivant :

Méthode faisceau
Soit $u > 0$, $m \in]0,1[$ et un point de départ $x^0 \in \mathbb{R}^n$.
1. Obtenir $f(x^0)$ et $\xi^0 \in \partial f(x^0)$ par l'oracle.
2. Initialisation : $J_0 = \{0\}$, $\alpha_0^0 = 0$.
3. Pour $k = 0, 1, ...$
4. Résoudre le problème (3.9) ou son dual (3.4) pour obtenir v^k et d^k.
5. Si $v^k = 0$ alors STOP $\Longrightarrow$ x^k est minimum de f.
6. On pose $y^{k+1} = x^k + d^k$.
7. Obtenir $f(y^{k+1})$ et $\xi^{k+1} \in \partial f(y^{k+1})$ par l'oracle.
8. Si $f(y^{k+1}) < f(x^k) + mv^k$ alors {pas sérieux}
9. Poser $x^{k+1} = y^{k+1}$.
10. Poser $\alpha_j^{k+1} = \alpha_j^k + f(x^{k+1}) - f(x^k) - < \xi^j, d^k > \quad \forall j \in J_k$.
11. Poser $\alpha_{k+1}^{k+1} = 0$.
12. Sinon {pas nul}
13. Poser $x^{k+1} = x^k$.
14. Poser $\alpha_{k+1}^{k+1} = f(x^k) - f(y^{k+1}) + < \xi^{k+1}, d^k >$.
15. Fin si
16. Poser $J_{k+1} \subseteq J_k \cup \{k+1\}$ avec $k+1 \in J_{k+1}$.
17. Fin pour

3.3 Améliorations pratiques

Pour obtenir un algorithme efficace, il reste à discuter du choix du paramètre u et de la taille du faisceau. Faire varier intelligement u est essentiel pour que la méthode faisceau soit performante. En effet, si u est grand, la solution en d du problème (3.9) sera assez petite et si u est petit, la solution en d du problème (3.9) sera plutôt grande. Quelques heuristiques ont été inventées sur base de cette remarque. Quand trop de pas nuls sont pris, u doit être augmenté pour forcer le point de test à être plus proche de x^k et ainsi obtenir un modèle plus précis. D'autre part, quand trop de petits pas sérieux sont pris, u doit diminuer pour permettre au point de test de se

trouver à une plus grande distance de x^k pour ainsi espérer prendre de plus grands pas.
Par ailleurs, remarquons que rien n'est mentionné dans la méthode faisceau pour éviter que la taille du faisceau $|J_k|$ ne grandisse sans borne. De plus, la méthode peut ne pas converger si J_k est choisi trop petit. La solution proposée par Kiwiel consiste à ajouter au faisceau le sous-gradient agrégé $\bar{\xi}^k = \sum_{j \in J_k} \lambda_j^k \xi^j$ et $\bar{\alpha}^k = \sum_{j \in J_k} \lambda_j^k \alpha_j^k$. Plus précisément, Kiwiel a démontré la convergence de la méthode faisceau en incluant dans le faisceau (de l'étape $k+1$) ξ^{k+1}, $\bar{\xi}^k$ et tout sous-ensemble de sous-gradients présents dans le faisceau à l'étape k. On s'attend bien entendu à ce que des faisceaux riches en informations donnent lieu à des convergences plus rapides. Il faudra donc faire un compromis entre vitesse de convergence et utilisation de mémoire.

Chapitre 4

Méthodes du Point Proximal

Dans ce chapitre, nous examinons les méthodes faisceaux d'un point de vue différent. Nous construisons un algorithme qui généralise les méthodes faisceaux en différents points. Par exemple, il permet l'utilisation de modèles plus généraux que le modèle du plan sécant utilisé dans les méthodes faisceaux. De plus, cet algorithme est conceptuellement facile à généraliser pour qu'il tienne compte des erreurs produites lors de l'évaluation de la fonction objectif et de ses sous-gradients. Nous développerons une telle extension dans le chapitre suivant.
Jusqu'ici, notre objectif était de construire à chaque itération une direction de descente pour la fonction objectif convexe non différentiable f à partir d'informations collectées dans un voisinage de l'itéré. Dans ce chapitre, nous désirons construire une fonction convexe différentiable approximant la fonction f telle que les minimums de f et de son approximation coïncident. Puisque la fonction approximante, appelée régularisation de Moreau-Yosida de f, est convexe et différentiable, nous pourrons alors utiliser les méthodes d'optimisation classiques telles que la méthode du gradient, la méthode BFGS ...

4.1 Régularisation de Moreau-Yosida

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe. Considérons la perturbation quadratique $\tilde{f}_M : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par

$$\tilde{f}_M(x, y) := f(y) + \frac{1}{2}\|y - x\|_M^2$$

où M est une matrice symétrique définie positive. Pour $x \in \mathbb{R}^n$ fixé, la fonction $y \to \tilde{f}_M(x, y)$ est fortement convexe puisque f est convexe et la fonction

$y \to \frac{1}{2}\|y-x\|_M^2$ est fortement convexe ($\forall y \in \mathbb{R}^n \ \nabla_y^2 \frac{1}{2}\|y-x\|_M^2 = M$ où M est définie positive). Dès lors, pour x fixé, la fonction $y \to \tilde{f}_M(x,y)$ admet un unique minimum et la définition suivante a du sens.

Définition 4.1.1 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ convexe et $M \in S^n$ où $M > 0$. La fonction $f_M : \mathbb{R}^n \to \mathbb{R}$ définie par*

$$f_M(x) := \min_y \ \tilde{f}_M(x,y) \tag{4.1}$$

est la régularisation de Moreau-Yosida de f associée à la métrique M. L'unique minimum dans (4.1) noté $p_M^f(x)$ est appelé point proximal de x associé à f et M.

Remarque 4.1.1 *Le minimum dans (4.1) est obtenu lorsque $0 \in \partial_y \tilde{f}_M(x,y)$. Par la Proposition 2.1.4, il est clair que*

$$\begin{aligned} \partial_y \tilde{f}_M(x,y) &= \partial f(y) + \tfrac{1}{2}\partial \|y-x\|_M^2 \\ &= \partial f(y) + M(y-x). \end{aligned}$$

Puisque $y = p_M^f(x)$ est l'unique minimum, on conclut que $p_M^f(x)$ est l'unique point $y \in \mathbb{R}^n$ tel que

$$M(x-y) \in \partial f(y).$$

Commençons par établir les propriétés de la régularisation f_M.

Théorème 4.1.1 *La régularisation de Moreau-Yosida possède les propriétés suivantes :*

1. *f_M est convexe et minore f.*
2. *f_M est différentiable avec pour gradient*

$$\nabla f_M(x) = s_M^f(x) = M(x - p_M^f(x)) \in \partial f(p_M^f(x)). \tag{4.2}$$

3. *∇f_M est Lipschitz continu sur $\mathbb{R}^n$ de constante $\lambda_{\max}(M)$ i.e.*

$$\forall x, z \in \mathbb{R}^n \quad \|\nabla f_M(x) - \nabla f_M(z)\| \leq \lambda_{\max}(M)\|x - z\|.$$

Preuve :

1. Soient $x_1, x_2 \in \mathbb{R}^n$ et $t \in]0,1[$. Posons $y_1 = p_M^f(x_1)$ et $y_2 = p_M^f(x_2)$. On peut écrire successivement

$$
\begin{aligned}
& t f_M(x_1) + (1-t) f_M(x_2) \\
& \quad = t f(y_1) + (1-t) f(y_2) + \tfrac{t}{2}\|y_1 - x_1\|_M^2 + \tfrac{(1-t)}{2}\|y_2 - x_2\|_M^2 \\
& \quad \geq f(ty_1 + (1-t)y_2) + \tfrac{t}{2}\|y_1 - x_1\|_M^2 + \tfrac{(1-t)}{2}\|y_2 - x_2\|_M^2 \\
& \quad \geq f(ty_1 + (1-t)y_2) + \tfrac{1}{2}\|t(y_1 - x_1) + (1-t)(y_2 - x_2)\|_M^2 \\
& \quad = f(\bar{x}) + \tfrac{1}{2}\|\bar{x} - tx_1 - (1-t)x_2\|_M^2 \\
& \quad \geq \min_{y \in \mathbb{R}^n}\{f(y) + \tfrac{1}{2}\|y - tx_1 - (1-t)x_2\|_M^2\} \\
& \quad = f_M(tx_1 + (1-t)x_2),
\end{aligned}
$$

   d'où f_M est convexe. De plus,

$$\forall x \in \mathbb{R}^n \quad f_M(x) = \min_y \tilde{f}_M(x,y) \leq \tilde{f}_M(x,x) = f(x),$$

   d'où f_M minore f.

2. Il suffit de prouver que $f'_M(x,d) = (M(x - p_M^f(x)))^T d \quad \forall d \in R^n$ car alors

$$
\begin{aligned}
\partial f_M(x) &= \{\xi \in \mathbb{R}^n \mid < \xi, d > \leq f'(x,d) \quad \forall d \in \mathbb{R}^n\} \\
&= \{\xi \in \mathbb{R}^n \mid < \xi, d > \leq < M(x - p_M^f(x)), d > \quad \forall d \in \mathbb{R}^n\} \\
&= \{\xi \in \mathbb{R}^n \mid < \xi - M(x - p_M^f(x)), d > \leq 0 \quad \forall d \in \mathbb{R}^n\} \\
&= \{M(x - p_M^f(x))\},
\end{aligned}
$$

   et par la Proposition 2.1.2, on peut conclure que f_M est différentiable et que $\nabla f_M(x) = M(x - p_M^f(x))$. De plus, par la Remarque 4.1.1, $M(x - p_M^f(x)) \in \partial f(p_M^f(x))$.

Soient $d \in \mathbb{R}^n$, $d \neq 0$ et $t > 0$. Alors

$$\begin{aligned}\frac{f_M(x+td)-f_M(x)}{t} &\leq \frac{f(p_M^f(x))+(1/2)\|p_M^f(x)-x-td\|_M^2-f_M(x)}{t} \\ &= \frac{f(p_M^f(x))+(1/2)\|p_M^f(x)-x\|_M^2+(t^2/2)\|d\|_M^2}{t} \\ &\quad + \frac{t<M(x-p_M^f(x)),d>-f_M(x)}{t} \\ &= (t/2)\|d\|_M^2 + < M(x-p_M^f(x)), d >\end{aligned}$$

d'où $f'_M(x,d) \leq < M(x-p_M^f(x)), d > \ \forall d \in \mathbb{R}^n$. Puisque $f'_M(x,.)$ est convexe (admis), nous avons

$$0 = f'_M(x, \frac{1}{2}d + \frac{1}{2}(-d)) \leq \frac{1}{2}f'_M(x,d) + \frac{1}{2}f'_M(x,-d)$$

et

$$\begin{aligned}f'_M(x,d) &\geq -f'_M(x,-d) \\ &\geq - < M(x-p_M^f(x)), -d > \\ &= < M(x-p_M^f(x)), d > .\end{aligned}$$

Par conséquent, $f'_M(x,d) = (M(x-p_M^f(x)))^T d$ pour tout $d \in \mathbb{R}^n$.

3. Soient $x, z \in \mathbb{R}^n$. Par la deuxième partie du théorème, nous avons

$$\begin{aligned}\nabla f_M(x) - \nabla f_M(z) &= M(x-p_M^f(x)) - M(z-p_M^f(z)) \\ &= M(x-z) - M(p_M^f(x) - p_M^f(z)).\end{aligned}$$

En prenant le produit scalaire des membres de gauche et de droite avec $M^{-1}(\nabla f_M(x) - \nabla f_M(z))$, on obtient

$$\begin{aligned}\|\nabla f_M(x) - \nabla f_M(z)\|_{M^{-1}}^2 &= < \nabla f_M(x) - \nabla f_M(z), x - z > \\ &\quad - < \nabla f_M(x) - \nabla f_M(z), p_M^f(x) - p_M^f(z) > .\end{aligned}$$

Puisque $\nabla f_M(x) \in \partial f(p_M^f(x))$, $\nabla f_M(z) \in \partial f(p_M^f(z))$ et l'opérateur ∂f est monotone (par la Proposition 2.1.6), nous avons

$$< \nabla f_M(x) - \nabla f_M(z), p_M^f(x) - p_M^f(z) > \geq 0.$$

En utilisant l'expression précédente et l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned}\|\nabla f_M(x) - \nabla f_M(z)\|^2_{M^{-1}} &\leq < \nabla f_M(x) - \nabla f_M(z), x - z > \\ &\leq \|\nabla f_M(x) - \nabla f_M(z)\| \, \|x - z\|.\end{aligned}$$

On obtient finalement que

$$\lambda_{\min}(M^{-1})\|\nabla f_M(x) - \nabla f_M(z)\| \leq \|x - z\|.$$

□

La régularisation de Moreau-Yosida est en fait une approximation de classe C^1 qui peut être rendue arbitrairement proche de la fonction originale non différentiable. Le lemme suivant en témoigne.

Lemme 4.1.1 *Si $\lambda_{\min}(M) \to \infty$ alors $f_M(x) \to f(x)$ et $p^f_M(x) \to x$ pour tout $x \in \mathbb{R}^n$. De plus, pour $M > 0$ fixée, si $\mu \nearrow \infty$ alors*

$$\begin{aligned} f_{\mu M}(x) &\nearrow f(x), \\ \|p^f_{\mu M}(x) - x\| &\searrow 0. \end{aligned}$$

Il est maintenant temps de démontrer l'équivalence entre minimiser f_M et minimiser f. Pour ce faire, nous aurons besoin du lemme suivant.

Lemme 4.1.2 *Pour tout $x \in \mathbb{R}^n$,*

$$f(p^f_M(x)) \leq f(x) - \|s^f_M(x)\|^2_{M^{-1}}. \tag{4.3}$$

Preuve : Par (4.2), $s^f_M(x) = M(x - p^f_M(x)) \in \partial f(p^f_M(x))$ d'où

$$f(x) \geq f(p^f_M(x)) + < s^f_M(x), x - p^f_M(x) > .$$

Or, $x - p^f_M(x) = M^{-1}s^f_M(x)$. On obtient alors

$$\begin{aligned} f(x) &\geq f(p^f_M(x)) + < s^f_M(x), M^{-1}s^f_M(x) > \\ &= f(p^f_M(x)) + \|s^f_M(x)\|^2_{M^{-1}}. \end{aligned}$$

□

Théorème 4.1.2 *Les assertions suivantes sont équivalentes :*

1. $\bar{x}$ *minimise* f ;
2. $p_M^f(\bar{x}) = \bar{x}$;
3. $s_M^f(\bar{x}) = 0$;
4. $\bar{x}$ *minimise* f_M ;
5. $f(p_M^f(\bar{x})) = f(\bar{x})$;
6. $f_M(\bar{x}) = f(\bar{x})$.

Preuve :
$1 \Rightarrow 2$. Si $\bar{x}$ minimise f, alors $y = \bar{x}$ minimise la fonction $y \to \tilde{f}_M(\bar{x}, y) = f(y) + \frac{1}{2}\|y - \bar{x}\|_M^2$ donc $\bar{x} = p_M^f(\bar{x})$ car $p_M^f(\bar{x})$ est l'unique minimum de $y \to \tilde{f}_M(\bar{x}, y)$.

$2 \Leftrightarrow 3$ car $s_M^f(\bar{x}) = M(\bar{x} - p_M^f(\bar{x}))$.

$3 \Leftrightarrow 4$ car $\nabla f_M(\bar{x}) = s_M^f(\bar{x})$ et f_M est convexe.

$4 \Rightarrow 5$ car $4 \Rightarrow 3 \Rightarrow 2 \Rightarrow 5$.

$5 \Rightarrow 4$ car $5 \Rightarrow 3 \Rightarrow 4$. En effet, par le lemme 4.1.2 , $\|s_M^f(\bar{x})\|_{M^{-1}}^2 \leq 0$ d'où $s_M^f(\bar{x}) = 0$.

$5 \Rightarrow 6$ car $f_M(\bar{x}) = f(p_M^f(\bar{x})) + \frac{1}{2}\|p_M^f(\bar{x}) - \bar{x}\|_M^2$. Puisque $5 \Rightarrow 2$, on conclut que $f_M(\bar{x}) = f(p_M^f(\bar{x})) = f(\bar{x})$.

$6 \Rightarrow 1.$

$$\begin{aligned} f(\bar{x}) &= f_M(\bar{x}) \\ &= f(p_M^f(\bar{x})) + \tfrac{1}{2}\|p_M^f(\bar{x}) - \bar{x}\|_M^2 \\ &\leq f(\bar{x}) - \|s_M^f\|_{M^{-1}}^2 + \tfrac{1}{2}\|p_M^f(\bar{x}) - \bar{x}\|_M^2 \quad \text{(par (4.3))} \\ &= f(\bar{x}) - < M^{-1}M(\bar{x} - p_M^f(\bar{x})), M(\bar{x} - p_M^f(\bar{x})) > \\ &\quad + \tfrac{1}{2}\|p_M^f(\bar{x}) - \bar{x}\|_M^2 \\ &= f(\bar{x}) - \|p_M^f(\bar{x}) - \bar{x}\|_M^2 + \tfrac{1}{2}\|p_M^f(\bar{x}) - \bar{x}\|_M^2 \\ &= f(\bar{x}) - \tfrac{1}{2}\|p_M^f(\bar{x}) - \bar{x}\|_M^2 \end{aligned}$$

d'où 2 et 3. Par conséquent, $0 = s_M^f(\bar{x}) \in \partial f(p_M^f(\bar{x})) = \partial f(\bar{x})$.

□

On remarque suite à ce théorème que minimiser f_M est équivalent à minimiser f dans le sens où elles ont les mêmes valeurs minimales et les mêmes minimums. Par ailleurs, f_M possède un gradient Lipschitz continu sur $\mathbb{R}^n$. La fonction f_M est donc une candidate idéale pour une méthode d'optimisation différentiable classique telle que la méthode BFGS. Malheureusement, une telle approche présente un inconvénient de taille. En effet, l'évaluation de $f_M(x)$ et de son gradient est elle-même un problème d'optimisation aussi difficile que le problème de départ (excepté le faible gain de la forte convexité). Pour commencer, nous allons ignorer cette difficulté afin de décrire le concept de l'algorithme du point proximal. Cet algorithme servira ensuite de fondations pour une version plus implémentable où des erreurs dans l'évaluation de $f_M(x)$ et de $p_M^f(x)$ seront tolérées.

4.2 Algorithme du point proximal

Par le Théorème 4.1.2, minimiser f est équivalent à trouver un point fixe à l'opérateur prox p_M^f d'où l'itération $x^{k+1} = p_M^f(x^k)$ pour trouver un minimum de f. Cet algorithme est connu sous le nom d'algorithme du Point Proximal.

Algorithme du Point Proximal (métrique fixée)
1. Choisir $x^0 \in \mathbb{R}^n$ et $M > 0$. Poser $k = 0$.
2. Calculer $x^{k+1} = p_M^f(x^k)$ en résolvant le problème

$$\min_{y \in \mathbb{R}^n} \{f(y) + \frac{1}{2}\|y - x^k\|_M^2\}.$$

3. Si $x^{k+1} = x^k$ alors STOP $\implies$ x^{k+1} est minimum de f.
4. Remplacer k par $k+1$ et retourner à l'étape 2.

On peut démontrer que pour l'algorithme du Point Proximal, x^k converge vers un minimum si il en existe. La convergence de l'algorithme n'est pas démontrée car ce dernier constitue une base théorique pour des versions plus implémentables où les preuves seront décrites dans les détails.

Remarque 4.2.1 *En remplaçant $p_M^f(x)$ par $x - M^{-1}s_M^f(x)$ (voir (4.2)), l'algorithme du point proximal peut être réinterprété tel une méthode du gradient préconditionné appliquée à f_M*

$$x^{k+1} = x^k - M^{-1}\nabla f_M(x^k),$$

ou tel une méthode du sous-gradient préconditionné appliquée à f

$$x^{k+1} = x^k - M^{-1}\xi^k, \quad \xi^k = s_M^f(x^k) \in \partial f(p_M^f(x^k)) = \partial f(x^{k+1}),$$

où le sous-gradient utilisé à l'itération k est un sous-gradient du point d'itération $k+1$.

4.3 Algorithme du point proximal à métrique variable

L'algorithme du point proximal fonctionne quelque soit le choix de la matrice symétrique définie positive M dans la régularisation de Moreau-Yosida d'où la question du choix de M pour obtenir une convergence rapide. Nous étions déjà confrontés à une telle question concernant le paramètre u dans la méthode faisceau. Considérons la taille de M en terme de valeurs propres pour comprendre l'intérêt d'un choix judicieux de ce paramètre.

De grandes valeurs propres de M produisent moins de lissage, moins de "smoothing". En effet, selon le Lemme 4.1.1, plus $\lambda_{\min}(M)$ est grande et plus les distances $|f(x) - f_M(x)|$ et $\|x - p_M^f(x)\|$ sont petites ce qui signifie que $f_M(x)$ et $p_M^f(x)$ sont plus faciles à évaluer. Cependant, un plus grand nombre de pas dans l'algorithme du point proximal sont nécessaires pour converger vers un point ε-optimal.

De petites valeurs propres de M produisent plus de "smoothing" et donc davantage de travail dans l'évaluation de $f_M(x)$ et de $p_M^f(x)$. Cependant, un plus petit nombre de pas dans l'algorithme du point proximal sont nécessaires pour converger vers un point ε-optimal.

Le choix de la métrique M s'apparente donc à un choix entre la quantité de travail d'un pas de l'algorithme du point proximal et le nombre de pas nécessaire pour obtenir un point ε-optimal. Une métrique "optimale" en terme de travail total nécessaire est donc difficile à définir mais il est clair qu'elle doit dépendre à la fois du comportement de la fonction objectif et de la nature du critère d'arrêt pour l'approximation de $p_M^f(x)$. C'est pour cette raison que l'on utilise une métrique variable M_k qui sera ajustée par l'algorithme à partir d'informations (collectées pendant le processus d'optimisation) au sujet de la fonction et ce, dans le but d'accélérer la convergence et au prix de l'interprétation de Moreau-Yosida qui disparaît. L'algorithme de la section 4.2 devient alors :

Algorithme du Point Proximal (métrique variable)
1. Choisir $x^0 \in \mathbb{R}^n$ et $M_0 > 0$. Poser $k = 0$.
2. Calculer $x^{k+1} = p_{M_k}^f(x^k)$ en résolvant le problème

$$\min_{y \in \mathbb{R}^n} \{f(y) + \frac{1}{2}\|y - x^k\|_{M_k}^2\}.$$

3. Si $x^{k+1} = x^k$ alors STOP $\implies$ x^{k+1} est minimum de f.
4. Choisir $M_{k+1} > 0$, remplacer k par $k+1$ et retourner à l'étape 2.

4.4 Méthode du point proximal approximé

Dans cette section, nous admettons que l'algorithme du point proximal pur est irréalisable en pratique car chaque itération consiste en un problème

d'optimisation qui doit être résolu itérativement. Pour surmonter cet obstacle, nous utilisons une condition d'arrêt afin d'obtenir une terminaison finie de cette boucle "intérieure". L'algorithme en résultant doit ensuite être analysé afin de s'assurer que l'utilisation de la condition d'arrêt ne détruise pas la convergence de la boucle "extérieure". Commençons par étudier un critère d'arrêt conceptuel qui ne peut être implémenté pour ensuite déduire un critère d'arrêt suffisant qui peut l'être.

4.4.1 Condition d'arrêt conceptuelle

A l'itération k, nous avons le point x^k et une matrice symétrique définie positive M_k jouant le rôle de la métrique qui rappellons le peut varier d'une itération extérieure à l'autre. La boucle intérieure qui calcule $f_{M_k}(x^k)$ et $p^f_{M_k}(x^k)$ s'arrêtera en un point x^{k+1} satisfaisant un critère d'arrêt. Considérons le critère d'arrêt suivant

$$f(x^{k+1}) \leq f(x^k) - m\delta^k \tag{4.4}$$

où $m \in (0,1)$ est une constante et

$$\delta^k := f(x^k) - f_{M_k}(x^k). \tag{4.5}$$

Notons que $\delta^k \geq 0$ avec égalité si et seulement si x^k minimise f par le Théorème 4.1.1 (1) et le Théorème 4.1.2 (6). Cette condition ne peut bien entendu pas être testée directement car $f_{M_k}(x^k)$ n'est pas calculable. Plus tard, le Théorème 4.4.3 décrira une condition d'arrêt implémentable qui entraînera (4.4) mais pour l'instant, nous considérons directement (4.4). Les deux théorèmes suivants nous montrent que le critère d'arrêt (4.4) entraîne une terminaison finie de la boucle intérieure et la convergence de la boucle extérieure.

Théorème 4.4.1 *Soient M^k et x^k qui n'est pas minimum. Soit $\{y^j\} \subset \mathbb{R}^n$ une suite qui converge vers $p^f_{M_k}(x^k)$. Alors, il existe un $J \in \mathbb{N}$ tel que $x^{k+1} = y^J$ satisfait la condition (4.4).*

Preuve : Par (4.5), nous avons

$$\begin{aligned}\delta^k &= f(x^k) - f_{M_k}(x^k) \\ &= f(x^k) - f(p^f_{M_k}(x^k)) - \tfrac{1}{2}\|p^f_{M_k}(x^k) - x^k\|^2_{M_k} \qquad (4.6) \\ &\leq f(x^k) - f(p^f_{M_k}(x^k)).\end{aligned}$$

Puisque $y^j \to p^f_{M_k}(x^k)$ et f est continue, on sait que $f(y^j) \to f(p^f_{M_k}(x^k))$ d'où il existe un $J \in \mathbb{N}$ tel que

$$|f(y^J) - f(p^f_{M_k}(x^k))| \leq (1-m)\delta^k$$

ce qui entraîne que

$$-f(p^f_{M_k}(x^k)) \leq (1-m)\delta^k - f(y^J).$$

Combiné à (4.6), on obtient

$$\delta^k \leq f(x^k) - f(y^J) + (1-m)\delta^k$$

ou encore

$$f(y^J) \leq f(x^k) - m\delta^k,$$

d'où $x^{k+1} = y^J$ satisfait la condition (4.4).

□

Théorème 4.4.2 *Soit la suite $\{M_k\} \subset S^n$ telle que $M_k > 0$ pour tout k et*

$$\sum_{k=0}^{\infty} \frac{1}{\lambda_{\max}(M_k)} = \infty. \qquad (4.7)$$

Supposons que la suite $\{x^k\}$ est bornée et qu'elle satisfait la condition (4.4) pour tout k. Alors, tout point d'accumulation de $\{x^k\}$ minimise f et $f(x^k) \to \min_x f(x)$.

Preuve : La suite $\{f(x^k)\}$ est décroissante car la suite $\{x^k\}$ satisfait la condition (4.4) et $\{f(x^k)\}$ est bornée car $\{x^k\}$ est bornée et f est continue. Par conséquent, $f(x^k) \to \bar{f}$ où $\bar{f} > -\infty$.
De nouveau par (4.4),

$$m\delta^k \leq f(x^k) - f(x^{k+1}).$$

Pour tout $K \geq 0$, on obtient

$$m \sum_{k=0}^{K} \delta^k \leq \sum_{k=0}^{K} (f(x^k) - f(x^{k+1})) = f(x^0) - f(x^{K+1}).$$

La suite des sommes partielles $\{\sum_{k=0}^{K} \delta^k\}_{K \in \mathbb{N}}$ est une suite croissante alors que le membre de droite tend vers $f(x^0) - f$ quand $K \to \infty$. Par conséquent, $\sum_k \delta^k$ converge et $\delta^k \to 0$.

En partant de la définition de δ^k, on obtient la suite d'inégalités suivantes :

$$\begin{aligned}
\delta^k &= f(x^k) - f_{M_k}(x^k) \\
&= f(x^k) - f(p^f_{M_k}(x^k)) - \tfrac{1}{2}\|p^f_{M_k}(x^k) - x^k\|^2_{M_k} \\
&\geq f(x^k) - f(x^k) + \|s^f_{M_k}(x^k)\|^2_{M_k^{-1}} - \tfrac{1}{2}\|p^f_{M_k}(x^k) - x^k\|^2_{M_k} \quad \text{(par (4.3))} \\
&= \tfrac{1}{2}\|s^f_{M_k}(x^k)\|^2_{M_k^{-1}} \quad \text{(par (4.2))} \\
&\geq \tfrac{1}{2\lambda_{\max}(M_k)}\|s^f_{M_k}(x^k)\|^2.
\end{aligned}$$

Puisque $\sum_k \delta^k$ converge, $\sum_k \frac{1}{2}\|s^f_{M_k}(x^k)\|^2/\lambda_{\max}(M_k)$ converge. Par (4.7), on conclut que $\{s^f_{M_k}(x^k)\}$ doit avoir 0 comme point d'accumulation. Puisque $\{x^k\}$ est bornée, on peut en extraire une sous-suite indicée par $I \subset \mathbb{N}$ telle que $\lim_{k \in I} x^k = \bar{x}$ et $\lim_{k \in I} s^f_{M_k}(x^k) = 0$. On peut alors utiliser la formule du transfert pour transformer le sous-gradient $s^f_{M_k}(x^k) \in \partial f(p^f_{M_k}(x^k))$ en un ε^k-sous-gradient de f en x^k où

$$\begin{aligned}
\varepsilon^k &= f(x^k) - f(p^f_{M_k}(x^k)) - < s^f_{M_k}(x^k), x^k - p^f_{M_k}(x^k) > \\
&= f(x^k) - f(p^f_{M_k}(x^k)) - \|x^k - p^f_{M_k}(x^k)\|^2_{M_k} \\
&= f(x^k) - f_{M_k}(x^k) - \tfrac{1}{2}\|x^k - p^f_{M_k}(x^k)\|^2_{M_k} \\
&= \delta^k - \tfrac{1}{2}\|x^k - p^f_{M_k}(x^k)\|^2_{M_k}.
\end{aligned}$$

D'où $\delta^k \to 0$ implique $\varepsilon^k \to 0$. Par la Proposition 2.3.4, on peut alors conclure que $\bar{x}$ minimise f i.e. $f(\bar{x}) = \min_x f(x)$. Puisque $\bar{f}$ est l'unique point d'accumulation de $\{f(x^k)\}$ et $\lim_{k\in I} f(x^k) = \min_x f(x)$, $\bar{f}$ doit être égal à $\min_x f(x)$ et tout autre point d'accumulation de $\{x^k\}$ doit aussi minimiser f.

□

L'hypothèse (4.7) du Théorème 4.4.2 mérite quelques commentaires. La condition (4.7) sur la suite $\{M_k\}$ signifie grossièrement que M_k peut grandir jusque l'infini mais pas "trop rapidemment". Remarquons que le théorème est formulé de telle façon que la suite $\{M_k\}$ est spécifiée a priori mais, en fait, chaque matrice M_k peut être générée a postériori par la boucle intérieure pour peu que toute la suite obéisse à (4.7).

4.4.2 Condition d'arrêt pratique

Cette section présente une condition implémentable suffisante pour la condition d'arrêt conceptuelle (4.4). Elle est basée sur le remplacement de f par un modèle minorant φ pour lequel le calcul du point proximal est direct.

Théorème 4.4.3 *Soient x^k et M_k. Soit $\varphi : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe qui minore f. Définissons $\pi := p^{\varphi}_{M_k}(x^k)$. Alors,*

$$f(\pi) \leq f(x^k) - m(f(x^k) - \varphi(\pi)) \tag{4.8}$$

implique la condition (4.4) avec $x^{k+1} = \pi$. De plus, si x^k ne minimise pas f alors il existe $\rho^k > 0$ indépendant de φ et de π tel que $f(\pi) - \varphi(\pi) \leq \rho^k$ implique (4.8).

Preuve :

1. Pour la première conclusion, il suffit de prouver que $f_{M_k}(x^k) \geq \varphi(\pi)$ car dans ce cas, (4.8) implique

$$f(\pi) \leq f(x^k) - m(\underbrace{f(x^k) - f_{M_k}(x^k)}_{\delta^k})$$

qui implique à son tour (4.4) avec $x^{k+1} = \pi$.
Puisque φ minore f, il suit que φ_{M_k} minore f_{M_k} d'où

$$f_{M_k}(x^k) \geq \varphi_{M_k}(x^k) = \varphi(\pi) + \frac{1}{2}\|\pi - x^k\|^2_{M_k} \geq \varphi(\pi). \tag{4.9}$$

2. Soustrayons $\varphi(\pi)$ des deux membres de l'inégalité (4.8) pour obtenir la condition équivalente

$$f(\pi) - \varphi(\pi) \leq (1-m)(f(x^k) - \varphi(\pi)). \tag{4.10}$$

Puisque $\delta^k = f(x^k) - f_{M_k}(x^k) \leq f(x^k) - \varphi(\pi)$, posons $\rho^k = (1-m)\delta^k$ pour montrer que $f(\pi) - \varphi(\pi) \leq \rho^k$ implique (4.10) et donc (4.8). Finalement, $\rho^k > 0$ car δ^k est strictement positif si x^k ne minimise pas f.

□

Remarque 4.4.1 *La preuve du Théorème 4.4.3 ne nécessite pas vraiment que φ minore f. La preuve a uniquement besoin de l'inégalité $\varphi(p^f_{M_k}(x^k)) \leq f(p^f_{M_k}(x^k))$ afin d'obtenir $f_{M_k}(x^k) \geq \varphi_{M_k}(x^k)$ dans (4.9). Mais puisque $p^f_{M_k}(x^k)$ est inconnu, il est plus sûr d'utiliser un modèle minorant f globalement.*

Les deux affirmations du Théorème 4.4.3 nous permettent de conclure que si l'on dispose d'un algorithme construisant un modèle φ suffisamment proche de f, alors l'utilisation du point proximal de φ suffit pour obtenir la convergence décrite dans le Théorème 4.4.2. De plus, le modèle φ doit seulement être proche de f au point proximal π.
Le rapport entre la méthode du point proximal approximé et la méthode faisceau peut maintenant être décrit. Pour que la méthode du point proximal approximé ressemble à la méthode faisceau, la métrique utilisée dans la régularisation de Moreau-Yosida doit être de la forme $M_k \equiv uI$ et la fonction modèle doit correspondre au modèle du plan sécant construit à partir des linéarisations de f aux différents points proximaux approximés. Le calcul de $p^{\varphi}_{uI}(x^k)$ correspond à la résolution du sous-problème de direction de recherche pour calculer le nouveau point test. Les itérations extérieures de la méthode du point proximal approximé correspondent évidemment aux pas de descente de la méthode faisceau. Les itérations intérieures sont les pas nuls qui améliorent le modèle en ajoutant de nouveaux sous-gradients jusqu'à ce que la condition d'arrêt (4.8), identique au test de descente dans la méthode faisceau, soit vérifiée.
La méthode du point proximal approximé est évidemment plus avantageuse que la méthode faisceau puisqu'elle généralise en quelque sorte la méthode faisceau. En effet, la méthode du point proximal approximé permet l'utilisation d'une métrique arbitraire M_k, d'un modèle convexe arbitraire φ et d'une boucle intérieure arbitraire pour améliorer le modèle.

4.4.3 Amélioration du modèle

L'algorithme de recherche d'un modèle φ tel que $f(\pi)-\varphi(\pi)$ soit suffisamment petit détermine les itérations intérieures, indicées par j, de la méthode du point proximal approximé. Pour simplifier les notations, on supprime l'indice k de la boucle extérieure des quantités comme x^k et M_k puisque toutes les actions prennent place à l'intérieur d'une itération extérieure. La notation du point proximal peut alors être simplifiée par :

$$y^j := p_M^{\varphi^j}(x) \qquad g^j := s_M^{\varphi^j}(x) = M(x - y^j) \in \partial\varphi^j(y^j). \tag{4.11}$$

Finalement, pour toute fonction convexe ψ^j, la notation de la pertubation quadratique est simplifiée par :

$$\tilde{\psi}^j(y) := \tilde{\psi}_M^j(x, y) = \psi^j(y) + \frac{1}{2}\|y - x\|_M^2.$$

Nous allons construire une suite de modèles $\{\varphi^j\}$ de tel façon que $f(y^j) - \varphi^j(y^j) \to 0$. Par conséquent, le Théorème 4.4.3 nous garantit que la condition d'arrêt (4.8) sera satisfaite en un nombre fini d'étapes si x ne minimise pas f. Si x minimise f alors nous verrons que $y^j \to x$. L'algorithme utilisera la fonction agrégation $l^j : \mathbb{R}^n \to \mathbb{R}$ définie par

$$l^j(y) := \varphi^j(y^j) + < g^j, y - y^j > . \tag{4.12}$$

Puisque $g^j \in \partial\varphi^j(y^j)$, l^j minore φ^j. Le lemme suivant nous montre que la perturbation quadratique de la fonction agrégation l^j est centrée en y^j.

Lemme 4.4.1 *L'égalité suivante est vérifiée :*

$$\tilde{l}^j(y) = \tilde{l}^j(y^j) + \frac{1}{2}\|y - y^j\|_M^2. \tag{4.13}$$

Preuve :

$$\begin{aligned}
\tilde{l}^j(y) &= \varphi^j(y^j) + < g^j, y - y^j > + \tfrac{1}{2}\|y - x\|_M^2 \\
&= \varphi^j(y^j) - < M(y^j - x), y - y^j > + \tfrac{1}{2}\|(y^j - x) + (y - y^j)\|_M^2 \\
&= \varphi^j(y^j) + \tfrac{1}{2}\|y^j - x\|_M^2 + \tfrac{1}{2}\|y - y^j\|_M^2 \\
&= l^j(y^j) + \tfrac{1}{2}\|y^j - x\|_M^2 + \tfrac{1}{2}\|y - y^j\|_M^2 \\
&= \tilde{l}^j(y^j) + \tfrac{1}{2}\|y - y^j\|_M^2.
\end{aligned}$$

□

Nous imposons trois conditions sur les modèles φ^j pour tout j :

$$\varphi^{j+1}(y) \leq f(y) \quad \forall y \in \mathbb{R}^n, \tag{4.14}$$

$$\varphi^{j+1}(y) \geq f(y^j) + < s(y^j), y - y^j > \quad \forall y \in \mathbb{R}^n, \tag{4.15}$$

$$\varphi^{j+1}(y) \geq l^j(y) \quad \forall y \in \mathbb{R}^n, \tag{4.16}$$

où $s(y^j)$ est le sous-gradient de f en y^j retourné par l'oracle. La première condition est l'habituelle propriété de minorisation, les deux suivantes indiquent comment le nouveau modèle doit être généré une fois y^j calculé. Par exemple, les modèles

. **Choix maximal (modèle du plan sécant) :**
$\varphi^{j+1}(y) = \max \{ f(y^i) + < s(y^i), y - y^i > \mid i = 0, ..., j \}$.

. **Choix minimal :**
$\varphi^{j+1}(y) = \max \{ l^j(y), f(y^j) + < s(y^j), y - y^j > \}$.

. **Choix intermédiaire :**
$\varphi^{j+1}(y) = \max [\{l^j(y)\} \cup \{ f(y^i) + < s(y^i), y - y^i > \mid i \in I^j \}]$ où $I^j \subset \{1, ..., j\}$ avec $j \in I^j$.

vérifient les trois critères. Bien entendu, d'autres modèles sont possibles.
La matrice M définissant la métrique ne peut être totalement arbitraire. En effet, la suite $\{M_k\}$ doit satisfaire (4.7) pour qu'il y ait convergence. Il est donc raisonnable d'imposer une borne telle que

$$M \leq \nu I < \infty. \tag{4.17}$$

Par conséquent, le spectre de M est compris dans l'intervalle $(0, \nu] \subset (0, \infty)$. Avec ces définitions et contraintes, la convergence des itérations intérieures peut être démontrée.

Théorème 4.4.4 *Etant donné $x \in \mathbb{R}^n$. Supposons que les fonctions convexes $\{\varphi^j\}$, la métrique M et les points de test $\{y^j\}$ vérifient (4.11),(4.12) et (4.14)-(4.17). Alors,*

$$f(y^j) - \varphi^j(y^j) \to 0, \tag{4.18}$$

$$y^j \to p_M^f(x). \tag{4.19}$$

Preuve :

1. Considérons la suite d'inégalité suivante :

$$\begin{aligned}
f(x) &\geq \varphi^{j+1}(x) && \text{(par (4.14))}\\
&= \varphi^{j+1}(x) + \tfrac{1}{2}\|x - x\|_M^2\\
&= \tilde{\varphi}^{j+1}(x)\\
&\geq \tilde{\varphi}^{j+1}(y^{j+1}) && \text{(par (4.11))}\\
&= \varphi^{j+1}(y^{j+1}) + \tfrac{1}{2}\|y^{j+1} - x\|_M^2\\
&= l^{j+1}(y^{j+1}) + \tfrac{1}{2}\|y^{j+1} - x\|_M^2 && \text{(par (4.12))}\\
&= \tilde{l}^{j+1}(y^{j+1})\\
&\geq l^j(y^{j+1}) + \tfrac{1}{2}\|y^{j+1} - x\|_M^2 && \text{(par (4.16))}\\
&= \tilde{l}^j(y^{j+1})\\
&= \tilde{l}^j(y^j) + \tfrac{1}{2}\|y^{j+1} - y^j\|_M^2. && \text{(par (4.13))}
\end{aligned}$$

De ces relations, on peut déduire que

- la suite $\{\tilde{l}^j(y^j)\}$ est croissante et bornée supérieurement par $f(x)$ d'où $\{\tilde{l}^j(y^j)\}$ est convergente.

- $\underbrace{\tilde{l}^{j+1}(y^{j+1}) - \tilde{l}^j(y^j)}_{\longrightarrow 0} \geq \tfrac{1}{2}\|y^{j+1} - y^j\|_M^2 \geq 0$ d'où $y^{j+1} - y^j \to 0$.

2. Soit $y \in \mathbb{R}^n$. Puisque $f \geq \varphi^j \geq l^j$, nous avons que

$$\begin{aligned} f(y) + \tfrac{1}{2}\|y - x\|_M^2 &\geq l^j(y) + \tfrac{1}{2}\|y - x\|_M^2 \\ &= \tilde{l}^j(y) \\ &= \tilde{l}^j(y^j) + \tfrac{1}{2}\|y - y^j\|_M^2. \quad \text{(par (4.13))} \end{aligned}$$

Puisque $\{\tilde{l}^j(y^j)\}$ est convergente, la suite $\{y - y^j\}$ doit être bornée. Par conséquent, la suite $\{y^j\}$ est bornée.

3. Les conditions sur les modèles impliquent que

$$\begin{aligned} f(y^{j+1}) - f(y^j) &\geq \varphi^{j+1}(y^{j+1}) - f(y^j) \quad \text{(par (4.14))} \\ &\geq < s(y^j), y^{j+1} - y^j > . \quad \text{(par (4.15))} \end{aligned}$$

Puisque $\{y^j\}$ est bornée et que la fonction f est Lipschitz continue sur les ensembles bornés, nous obtenons

$$|f(y^{j+1}) - f(y^j)| \leq L\|y^{j+1} - y^j\|,$$

d'où $f(y^{j+1}) - f(y^j) \to 0$ (car $y^{j+1} - y^j \to 0$).
D'autre part, la suite $\{s(y^j)\}$ est bornée car $\{y^j\}$ est bornée et le sous-différentiel est borné sur les ensembles bornés. Par conséquent,
$< s(y^j), y^{j+1} - y^j > \to 0$ car $y^{j+1} - y^j \to 0$.
On peut alors alors conclure que $\varphi^{j+1}(y^{j+1}) - f(y^j) \to 0$ et

$$\varphi^{j+1}(y^{j+1}) - f(y^{j+1}) = \varphi^{j+1}(y^{j+1}) - f(y^j) + f(y^j) - f(y^{j+1}) \to 0.$$

4. Puisque la suite $\{y^j\}$ est bornée, il suffit de prouver que tout point d'accumulation $\bar{y}$ de $\{y^j\}$ est égal à $p_M^f(x)$.
Supposons que $y^j \to \bar{y}$ pour $j \in K \subseteq \mathbb{N}$ (K infini). Alors, pour tout $j \in K$ et $y \in \mathbb{R}^n$, il suit de la définition du sous-différentiel que

$$f(y) \geq \varphi^j(y) \geq \varphi^j(y^j) + < g^j, y - y^j > . \qquad (4.20)$$

Puisque $\lim_{j\in K} f(y^j) = f(\bar{y})$ (continuité de f) et $\varphi^j(y^j) - f(y^j) \to 0$, nous avons que $\lim_{j\in K} \varphi^j(y^j) = f(\bar{y})$. De plus, $\lim_{j\in K} g^j = M(x - \bar{y})$. En passant à la limite pour $j \in K$ dans (4.20), nous obtenons pour tout y :

$$f(y) \geq f(\bar{y}) + < M(x - \bar{y}), y - \bar{y} >$$

i.e. $\bar{g} = M(x - \bar{y}) \in \partial f(\bar{y})$ d'où $\bar{y} = p_M^f(x)$ par la Remarque 4.1.1.

□

Corollaire 4.4.1 *Sous les hypothèses du Théorème 4.4.4, x minimise f si et seulement si $y^j \to x$.*

Preuve :
$\Rightarrow$: Si x minimise f alors, par le Théorème 4.1.2, $p_M^f(x) = x$ d'où $y^j \to p_M^f(x) = x$ par (4.19).

$\Leftarrow$: Si $y^j \to x$ alors, par (4.19) et par unicité de la limite, $p_M^f(x) = x$ d'où x minimise f par le Théorème 4.1.2.

□

4.4.4 Algorithme du point proximal approximé

Terminons par assembler les différentes pièces du puzzle pour obtenir l'algorithme du point proximal approximé. La convergence de l'algorithme est démontrée dans un unique théorème et un théorème supplémentaire nous indique qu'un point ε-stationnaire peut être obtenu en un temps fini.

Algorithme du point proximal approximé
Soient $\nu_{\max} > 0$, $\varepsilon_{\text{tol}} \geq 0$, $m \in (0,1)$ et un point de départ $x^0 \in \mathbb{R}^n$.
1. Pour $k = 0, 1, ...$
2. Choisir M_k tel que $0 < M_k \leq \nu_{\max} I$.
3. Choisir un modèle φ^1 minorant f tel que $\varphi^1(x^k) = f(x^k)$.
4. Poser $j = 0$.
5. Répéter
6. $j \leftarrow j + 1$.
7. Calculer $y^j = p_{M_k}^{\varphi^j}(x^k)$ et $g^j = s_{M_k}^{\varphi^j}(x^k)$.
8. Si $f(y^j) - \varphi^j(y^j) \leq \varepsilon_{\text{tol}}$ et $\|g^j\| \leq \varepsilon_{\text{tol}}$ alors
9. Poser $\bar{x} = y^j$ et STOP $\Rightarrow$ $\bar{x}$ est un point ε_{tol}-stationnaire.
10. Fin si
11. Choisir un modèle φ^{j+1} satisfaisant (4.14)-(4.16).
12. Jusqu'à ce que $f(y^j) \leq f(x^k) - m(f(x^k) - \varphi^j(y^j))$
13. Poser $x^{k+1} = y^j$.
14. Fin pour

Théorème 4.4.5 *Soit $f : I\!R^n \to I\!R$ convexe et soit la suite $\{x^k\}$ donnée par l'algorithme du point proximal approximé avec $\varepsilon_{tol} = 0$. Si la suite $\{x^k\}$ est finie et si x^K est le dernier élément de la suite, x^K minimise f. Sinon, si $\{x^k\}$ est bornée, ses points d'accumulations minimisent f et $f(x^k) \to \min_x f(x)$.*

Preuve :

1. Supposons que $\{x^k\}$ est finie et considérons la suite $\{y^j\}$ de l'itération extérieure K. Cette suite doit être infinie et nous avons $f(y^j) - \varphi^j(y^j) \to 0$ par le Théorème 4.4.4. Par le Théorème 4.4.3, x^K doit minimiser f car sinon, la condition d'arrêt en ligne 12 devrait être satisfaite pour un certain j ce qui est contraire au fait que $\{y^j\}$ est infinie.

2. Supposons que $\{x^k\}$ est infinie. Par le Théorème 4.4.3, la suite satisfait (4.4) et (4.7) est vérifiée car $M_k \leq \nu_{\max} I$ pour tout k. Il reste à appliquer le Théorème 4.4.2 pour compléter la preuve.

□

Rappelons qu'un point $\bar{x}$ est ε-stationnaire si il existe $\xi \in \partial_\varepsilon f(\bar{x})$ avec $\|\xi\| \leq \varepsilon$.

Théorème 4.4.6 *Soit $f : I\!R^n \to I\!R$ convexe et bornée inférieurement. Soit $\varepsilon_{tol} > 0$. Alors, l'algorithme du point proximal approximé se termine en un temps fini avec un point ε_{tol}-stationnaire $\bar{x}$.*

Preuve : Remarquons que

$$g^j \in \partial_{\eta_j} f(y^j) \text{ où } \eta_j := f(y^j) - \varphi^j(y^j). \tag{4.21}$$

En effet, $\forall\, y \in I\!R^n$

$$\begin{array}{lll} f(y) & \geq \varphi^j(y) & \text{(par 4.14))} \\ & \geq \varphi^j(y^j) + < g^j, y - y^j > & \text{(par 4.11))} \\ & = f(y^j) + < g^j, y - y^j > - \underbrace{(f(y^j) - \varphi^j(y^j))}_{\eta_j}. & \end{array}$$

Par conséquent, si l'algorithme se termine (à la ligne 8), l'expression (4.21) nous indique que $\bar{x} = y^j$ est nécessairement un point ε_{tol}-stationnaire. Il ne

nous reste donc plus qu'à prouver la terminaison de l'algorithme :

Procédons par contradiction, supposons que l'algorithme ne se termine pas. Tout comme dans le théorème précédent, il y a deux cas à considérer : $\{x^k\}$ est finie ou $\{x^k\}$ est infinie. Dans les deux cas, nous allons montrer que la condition d'arrêt à la ligne 8 sera finalement satisfaite en un certain point afin de contredire l'hypothèse de non terminaison.
Supposons que $\{x^k\}$ est finie et considérons la suite infinie $\{y^j\}$ (car l'algorithme ne se termine pas) de l'itération extérieure finale K. Par le Théorème 4.4.4, nous avons que $\eta_j \to 0$. D'autre part, par le Théorème 4.4.3, x^K doit minimiser f car sinon, la condition d'arrêt en ligne 12 devrait être satisfaite pour un certain j ce qui est contraire au fait que $\{y^j\}$ est infinie. De ce fait, le Corollaire 4.4.1 nous indique que $y^j \to x^K$ et implique que $g^j \to 0$ puisque $\|g^j\| = \|M_K(x^K - y^j)\| \leq \nu_{\max}\|x^K - y^j\|$. Par conséquent, les quantités η_j et $\|g^j\|$ deviennent finalement plus petites que ε_{tol} ce qui contredit l'hypothèse de non terminaison.
Supposons que $\{x^k\}$ est infinie et que $\bar{g}^k, \bar{\eta}^k$ et $\bar{\varphi}^k$ correspondent aux derniers g^j, η^j et φ^j de l'itération extérieure k. Puisque la condition (4.8) est satisfaite avec $\pi = x^{k+1}$, nous avons que

$$f(x^{k+1}) - f(x^k) \leq m(\bar{\varphi}^k(x^{k+1}) - f(x^k)) \leq -m\delta^k \leq 0$$

d'où $\{f(x^k)\}$ est une suite décroissante bornée inférieurement (car f bornée inférieurement). Par conséquent, $\{f(x^k)\}$ est convergente d'où

$$\bar{\varphi}^k(x^{k+1}) - f(x^k) \to 0. \tag{4.22}$$

L'expression suivante

$$\underbrace{f(x^k) - \bar{\varphi}^k(x^{k+1})}_{\to 0} \geq f(x^{k+1}) - \bar{\varphi}^k(x^{k+1}) = \bar{\eta}^k \geq 0$$

nous permet alors de conclure que $\bar{\eta}^k \to 0$. Par ailleurs, la suite d'inégalité

$$\begin{aligned} f(x^k) - \bar{\varphi}^k(x^{k+1}) &\geq \bar{\varphi}^k(x^k) - \bar{\varphi}^k(x^{k+1}) \\ &\geq < \bar{g}^k, x^k - x^{k+1} > \quad (\text{car } \bar{g}^k \in \partial\bar{\varphi}^k(x^{k+1})) \\ &= \|\bar{g}^k\|^2_{M_k^{-1}} \\ &\geq 0 \end{aligned}$$

combinée à (4.22) et $M_k^{-1} \geq 1/\nu_{\max} I > 0$ implique que $\bar{g}^k \to 0$. De nouveau, la condition d'arrêt doit finalement être satisfaite ce qui contredit l'hypothèse de non terminaison.

□

4.5 Résultats numériques

Nous nous proposons de traiter à l'aide de la méthode du point proximal approximé le problème test d'optimisation non différentiable appelé Max-quad. La fonction à minimiser est définie sur $\mathbb{R}^n$ et correspond au maximum de cinq fonctions quadratiques :

$$f_j(x) = x^T C^j x - d_j^T x,\ j = 1, ..., 5$$

où C^j est une matrice symétrique $n \times n$ définie par

$$C_{ik}^j = \exp(\frac{i}{k}) \cos(ik) \sin j,\ i < k \qquad C_{ii}^j = \frac{i}{n} |\sin j| + \sum_{i \neq k} |C_{ik}^j|$$

et d^j est un vecteur de $\mathbb{R}^n$ dont les composantes sont $d_i^j = \exp(i/j) \sin(ij)$. L'implémentation a été réalisée en langage MATLAB. Les principales caractéristiques de l'algorithme sont :

- le paramètre m est initialisé à 0.4,
- le point de départ est $x_0 = (1, ..., 1)$,
- le critère d'arrêt pour la boucle extérieure est $\|x^{k+1} - x^k\| \leq \eta$ où $\eta = 10^{-3}$,
- la condition d'arrêt de la boucle intérieure est supprimée,
- le faisceau est vidé après chaque pas sérieux,
- il permet de résoudre le problème via les modèles maximal et minimal,
- le nombre n de variables dans le problème à résoudre est fixé à 10.

Code MATLAB :

```
fip=fopen('résultat.doc','w');
n=10;m=5;
x0=ones(n,1);xold=x0 + ones(n,1);
[phi,s]= maxquadf(x0,n,m);
eta = 0.001;
A1= []; b1=[];
lambda = 25; sigma = 0.4;
H= zeros(n+1,n+1);
for j=1:n
   H(j,j)=lambda;
```

```
   lb(j)=-Inf;
   ub(j)=Inf;
end
lb(n+1) = -Inf; ub(n+1) = Inf;
lb=lb';ub=ub';
k=1; itkmax = 100; itimax = 200;
x=x0; v=0;
fprintf(fip,'Starting point : \n');
fprintf(fip,'%6.0f',x0);fprintf(fip,'\n');
fprintf(fip,'\n');
fprintf(fip,'Lower bounds : \n');
fprintf(fip,'%6.0f',lb);fprintf(fip,'\n');
fprintf(fip,'Upper bounds : \n');
fprintf(fip,'%6.0f',ub);fprintf(fip,'\n');
fprintf(fip,'\n');
fprintf(fip,'lambda = %6.4f \n',lambda);
fprintf(fip,'sigma = %6.4f \n',sigma);
fprintf(fip,'Stopping criterion: norm(xold -x) <= %7.6f \n',eta);
fprintf(fip,'\n');
fprintf(fip,'Initial value of phi : %8.4f',phi);
fprintf(fip,'\n');
fprintf(fip,'Iterations k phi i \n');
t=cputime;
while (k <= itkmax)& (norm(xold -x)> eta)
    xold=x;
    [x,ii] = bundle(H,x,v,lb,ub,A1,b1,sigma,k,n,m,itimax);
    [phi,s]= maxquadf(x,n,m);
    k
    phi
    ii
    fprintf(fip,'%6.0f \t %12.8f \t %6.0f\n',k,phi,ii);
    k=k+1;
end
e=cputime-t;
fprintf(fip,'Solution : \n ');
fprintf(fip,'%8.4f',x);
fprintf(fip,'\n');
fprintf(fip,'cputime (in seconds) : %8.4f',e);fprintf(fip,'\n');
[phi,s]= maxquadf(x,n,m);
fprintf(fip,'phi : %12.8f',phi);fprintf(fip,'\n');
```

```
Fsnorm=norm(s);
fprintf(fip,'Norm of s : %8.4f',Fsnorm);
fprintf(fip,'\n');
dist=norm(xold - x);
fprintf(fip,'Norm(xold-x) : %8.6f',dist);
fclose(fip);
```

function [x,ii]= bundle(H,x,v,lb,ub,A,b,sigma,k,n,m,itimax)

```
f=[-H(1:n,1:n)*x; 1];
i=1;
y=x; [phi,s]= maxquadf(y,n,m);
phik=phi;
A= [A; s' -1]; b=[b; s'*y - phi];
z = [x; v];
z = quadprog(H,f,A,b,[],[],lb,ub,z);
y=z; y(n+1)=[];
[phi,s]= maxquadf(y,n,m); v=z(n+1);
left = sigma*(phik - v);
right = phik - phi;

while (left > right)&(i<= itimax)

    % MAXIMAL CHOICE:

    A= [A ; s' -1];
    b= [b;s'*y - phi];

    % MINIMAL CHOICE:

    %A=[];
    %b=[];
    %A= [A; s' -1; (H(1:n,1:n)*(x-y))' -1];
    %b=[b;s'*y - phi;(H(1:n,1:n)*(x-y))'*y-v];

    z = quadprog(H,f,A,b,[],[],lb,ub,z);
    y=z; y(n+1)=[];
    [phi,s]= maxquadf(y,n,m); v=z(n+1);
    left = sigma*(phik - v);
```

```
    right = phik - phi;
    i=i+1;
end

x=y; ii=i+1;
```

function [F,G] = maxquadf(x,n,m)

```
for j=1:m
    for i=1:n
        c(i,j)=exp(i/j)*sin(i*j);
    end
end

for j=1:m
    for i=1:n
        for k=i+1:n
            q(i,k,j) = exp(i/k)*cos(i*k)*sin(j);
            q(k,i,j) = q(i,k,j);
        end
    end
end

for j=1:m
    for i=1:n
        q(i,i,j) = (i/n)*abs(sin(j));
        for k=1:n
            if k~=i
                q(i,i,j)=q(i,i,j)+abs(q(i,k,j));
            end
        end
    end
end
for j=1:m
    phi(j)= x'* q(:,:,j) * x - c(:,j)'* x;
end

[F,mi] = max(phi);mi1=mi(1);
G=2*q(:,:,mi1)*x - c(:,mi1);
```

L'algorithme (version modèle maximal) a été lancé avec différentes métriques (constantes pendant le processus d'optimisation) afin de mesurer l'influence de ce paramètre sur le comportement général de la méthode du point proximal approximé. Dans le tableau suivant, M est la métrique, k correspond au nombre de pas sérieux et μ correspond au nombre moyen de pas nuls par itération extérieure.

M	k	μ	Valeur optimale
diag(1,...,1)	15	55.8	-0.8414065
diag(25,...,25)	19	9.47	-0.8413951
diag(50,...,50)	29	7.27	-0.8412801
diag(75,...,75)	42	7.14	-0.8411583

Ces résultats numériques correspondent bien au comportement (en fonction du choix de la métrique) décrit à la section 4.3.
Notons que les modèles maximal et minimal ont tous les deux été testés. Le modèle maximal a bien entendu entraîné une convergence plus rapide et ce, au prix d'une utilisation de mémoire plus importante. Le choix entre ces deux modèles s'apparente donc à un choix entre vitesse de convergence et utilisation de mémoire. Un autre argument qui entre en ligne de compte quant au choix du modèle est la difficulté rencontrée lors de la résolution des sous-problèmes quadratiques. En effet, plus le modèle est riche et plus les sous-problèmes quadratiques risquent d'être difficiles à résoudre. Par conséquent, un bon compromis consiste à utiliser le modèle intermédiaire.

Chapitre 5

Méthode du Point Proximal Inexact

Toutes les méthodes développées jusqu'ici étaient basées sur l'Hypothèse 3.0.1 qui supposait l'existence d'un oracle. Cette hypothèse est sensée pour un grand nombre de fonctions mais il existe des fonctions qui ne peuvent être évaluées exactement et pour lesquelles aucun sous-gradient exact n'est disponible. Cependant, certaines de ces fonctions peuvent être évaluées avec n'importe quel degré de précision (dans les limites de l'arithmétique à précision finie) en utilisant une méthode itérative. Ce chapitre est une tentative de correction de la section 4.5 de [6], il consiste à adapter la méthode du point proximal approximé étudiée précédemment afin qu'elle puisse minimiser ce type de fonctions. A ce titre, observons ce qu'il arrive dans l'analyse de la section 4.4 lorsque la fonction objectif est évaluée inexactement. La valeur exacte de la fonction est utilisée par l'algorithme à deux endroits : la condition (4.15) dans la mise à jour du modèle et dans la condition d'arrêt de la boucle intérieure (4.8). Dans la section suivante, nous considérons uniquement une modification de (4.15) pour utiliser des valeurs approximées et des sous-gradients approximés de la fonction. Nous modifirons ensuite la condition (4.8).

5.1 Mise à jour du modèle modifiée

Commençons par remplacer (4.15) par

$$\varphi^{j+1}(y) \geq f(y^j) + < \hat{s}(y^j), y - y^j > -\varepsilon \quad \forall y \in \mathbb{R}^n, \tag{5.1}$$

où $\varepsilon \geq 0$ et $\hat{s}(y^j) \in \partial_\varepsilon f(y^j)$. Les résultats du Théorème 4.4.4 sont bien entendu pertubés suite à cette modification. En particulier, les points d'accumulation de $\{y^j\}$ ne sont plus des points proximaux mais des points ε-proximaux.

Définition 5.1.1 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ convexe, $x \in \mathbb{R}^n$, $M \in S^n$ définie positive et $\varepsilon \geq 0$. $\bar{y}$ est un point ε-proximal de x associé à f et M si $\bar{y}$ est un ε-minimum de la fonction $y \to \tilde{f}_M(x, y)$.*

Remarquons que l'on retrouve la définition classique du point proximal pour $\varepsilon = 0$.

Lemme 5.1.1 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ convexe, $x \in \mathbb{R}^n$, $M \in S^n$ définie positive et $\varepsilon \geq 0$. Si $\bar{y}$ satisfait*

$$M(x - \bar{y}) \in \partial_\varepsilon f(\bar{y}) \tag{5.2}$$

alors $\bar{y}$ est un point ε-proximal de x associé à f et M.

Preuve : Définissons $g : \mathbb{R}^n \to \mathbb{R}$ où $g(y) := \frac{1}{2}\|y - x\|_M^2$ et définissons $h : \mathbb{R}^n \to \mathbb{R}$ par

$$h(y) := \tilde{f}_M(x, y) = f(y) + g(y). \tag{5.3}$$

Alors, $\bar{y}$ est un point ε-proximal en x associé à f et M si et seulement si $0 \in \partial_\varepsilon h(\bar{y})$. Par la Proposition 2.3.5,

$$\begin{aligned} \partial_\varepsilon h(y) &\supset \partial_\varepsilon f(y) + \partial g(y) \\ &= \partial_\varepsilon f(y) + M(y - x), \end{aligned}$$

d'où (5.2) implique $0 \in \partial_\varepsilon h(\bar{y})$.

□

Nous pouvons maintenant présenter les versions perturbées du Théorème 4.4.4 et du Corollaire 4.4.1.

Théorème 5.1.1 *Soit $\varepsilon \geq 0$. Supposons que les hypothèses du Théorème 4.4.4 où la condition (4.15) est remplacée par (5.1) sont vérifiées. Alors,*

$$\limsup_{j\to\infty} (f(y^j) - \varphi^j(y^j)) \leq \varepsilon. \tag{5.4}$$

De plus, tout point d'accumulation $\bar{y}$ de la suite $\{y^j\}$ est un point ε-proximal de x associé à f et M.

Preuve :

1. Puisque l'argument (4.15) n'est pas utilisé dans les deux premières parties de la preuve du Théorème 4.4.4, nous pouvons donc de nouveau conclure que $y^{j+1} - y^j \to 0$ et que la suite $\{y^j\}$ est bornée.

2. Les conditions sur les modèles impliquent que

$$\begin{aligned} f(y^{j+1}) - f(y^j) &\geq \varphi^{j+1}(y^{j+1}) - f(y^j) \qquad \text{(par (4.14))} \\ &\geq < \hat{s}(y^j), y^{j+1} - y^j > -\varepsilon. \quad \text{(par (5.1))} \end{aligned}$$

   Puisque $\{y^j\}$ est bornée et que f est une fonction Lipschitz continue sur les ensembles bornés, nous obtenons

$$|f(y^{j+1}) - f(y^j)| \leq L\|y^{j+1} - y^j\|,$$

   d'où $f(y^{j+1}) - f(y^j) \to 0$ (car $y^{j+1} - y^j \to 0$).
   D'autre part, la suite $\{\hat{s}(y^j)\}$ est bornée car $\{y^j\}$ est bornée et l'ε-sous-différentiel est borné sur les ensembles bornés. Par conséquent,
   $< \hat{s}(y^j), y^{j+1} - y^j > \to 0$ car $y^{j+1} - y^j \to 0$.
   On peut alors conclure que $\liminf_{j\to\infty} (\varphi^{j+1}(y^{j+1}) - f(y^j)) \geq -\varepsilon$ et

$$\varphi^{j+1}(y^{j+1}) - f(y^{j+1}) = \varphi^{j+1}(y^{j+1}) - f(y^j) + \underbrace{f(y^j) - f(y^{j+1})}_{\to 0}$$

   d'où

$$\liminf_{j\to\infty} (\varphi^{j+1}(y^{j+1}) - f(y^{j+1})) \geq -\varepsilon$$

   ou encore

$$\limsup_{j\to\infty} (f(y^{j+1}) - \varphi^{j+1}(y^{j+1})) \leq \varepsilon.$$

3. Supposons que $y^j \to \bar{y}$ pour $j \in K \subseteq \mathbb{N}$ (K infini). Alors, pour tout $j \in K$ et $y \in \mathbb{R}^n$, il suit de la définition du sous-différentiel que

$$f(y) \geq \varphi^j(y) \geq \varphi^j(y^j) + < g^j, y - y^j > . \tag{5.5}$$

Puisque $\lim_{j\in K} f(y^j) = f(\bar{y})$ (continuité de f) et $\liminf_{j\to\infty} (\varphi^j(y^j) - f(y^j)) \geq -\varepsilon$, nous avons que $\liminf_{j\in K} \varphi^j(y^j) \geq f(\bar{y}) - \varepsilon$. De plus, $\lim_{j\in K} g^j = M(x - \bar{y})$. En passant à la limite pour $j \in K$ dans (5.5), nous obtenons pour tout y :

$$f(y) \geq f(\bar{y}) + < M(x - \bar{y}), y - \bar{y} > -\varepsilon \tag{5.6}$$

i.e. $\bar{g} = M(x - \bar{y}) \in \partial_\varepsilon f(\bar{y})$ ce qui implique par le Lemme 5.1.1 que $\bar{y}$ est un point ε-proximal de x associé à f et M.

□

Corollaire 5.1.1 *Supposons que les hypothèses du Théorème 5.1.1 sont vérifiées. Soit $\bar{y}$ un point d'accumulation de $\{y^j\}$. Si $\bar{y} = x$ alors x est un ε-minimum de f. Par contre, si x est un σ-minimum de f où $\sigma \geq 0$ alors,*

$$\|\bar{y} - x\| \leq \sqrt{\frac{\varepsilon + \sigma}{\lambda_{\min}(M)}}. \tag{5.7}$$

Preuve :

1. La première conclusion suit immédiatement de (5.6) avec $\bar{y} = x$.
2. x est un σ-minimum de f d'où

$$f(\bar{y}) - f(x) \geq -\sigma. \tag{5.8}$$

Remplaçons y par x dans (5.6) pour obtenir

$$\begin{aligned} f(\bar{y}) - f(x) &\leq < M(x - \bar{y}), \bar{y} - x > + \varepsilon \\ &= -\|\bar{y} - x\|_M^2 + \varepsilon. \end{aligned}$$

On peut alors conclure que $\lambda_{\min}(M)\|\bar{y} - x\|^2 \leq \|\bar{y} - x\|_M^2 \leq \varepsilon + \sigma$ d'où (5.7).

□

Tout comme pour le Théorème 4.4.4, la signification du Théorème 5.1.1 est en rapport avec la condition d'arrêt du Théorème 4.4.3. Si x ne minimise pas f et si

$$\varepsilon < \rho = (1-m)\delta, \tag{5.9}$$

alors le Théorème 4.4.3 nous indique que des erreurs d'évaluation de la fonction de moins de ε impliquent malgré tout (4.8) en un nombre fini d'itérations intérieures. Bien que la condition (5.9) ne peut être testée puisque δ est inconnu, on peut au moins conclure qu'il existe une tolérance positive indépendante du modèle qui est suffisante pour la terminaison finie du processus intérieur tant que x n'est pas minimum de f.

5.2 Condition d'arrêt modifiée

Le but de cette section est de construire une règle pratique pour choisir ε et de modifier la condition (4.8). Le paramètre ε doit être choisi de façon à ce que (4.8) soit vérifiée c'est pourquoi nous considérons la condition équivalente (4.10). Si

$$\varepsilon < (1-m)(f(x) - \varphi^j(y^j)) \quad \forall j, \tag{5.10}$$

alors le Théorème 5.1.1 nous dit que (4.10) est en fin de compte satisfaite. Pour maintenir (5.10), nous devons bien entendu diminuer ε à chaque fois que l'expression en question n'est pas vérifiée. Par conséquent, on permet à la tolérance de varier dans la boucle intérieure. Nous remplaçons donc (5.1) par

$$\varphi^{j+1}(y) \geq f(y^j) + < \hat{s}(y^j), y - y^j > -\varepsilon^j \quad \forall y \in \mathbb{R}^n \tag{5.11}$$

où

- la suite $\{\varepsilon^j\}$ est positive et décroissante,
- $\hat{s}(y^j) \in \partial_{\varepsilon^j} f(y^j)$.

Cependant, ce n'est pas suffisant de maintenir (5.10) car la condition d'arrêt (4.8) (ou (4.10)) nécessite les valeurs exactes $f(\pi)$ et$f(x)$. Bien que la condition d'arrêt est en fin de compte satisfaite, elle ne peut être testée. Le lemme suivant établit une condition testable utilisant les valeurs approximées de la fonction et qui implique la condition (4.8).

Lemme 5.2.1 *Soient $x \in \mathbb{R}^n$ et $M \in S^n$ définie positive. Supposons que $\varphi : \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe qui minore f. Posons $\pi := p_M^{\varphi}(x)$. Etant données les tolérances $\varepsilon \geq 0$ et $\varepsilon_x \geq 0$, les valeurs approximées de la fonction $\hat{f}$ et $\hat{f}_x$ satisfont*

$$\hat{f} \leq f(\pi) \leq \hat{f} + \varepsilon,$$
$$\hat{f}_x \leq f(x) \leq \hat{f}_x + \varepsilon_x.$$

Alors,

$$\hat{f} + \varepsilon \leq \hat{f}_x - m(\hat{f}_x + \varepsilon_x - \varphi(\pi)) \tag{5.12}$$

implique (4.4) avec $x^k = x$, $x^{k+1} = \pi$ et $M_k = M$.

Preuve : Nous obtenons successivement

$$\begin{aligned} f(\pi) &\leq \hat{f} + \varepsilon \\ &\leq \hat{f}_x - m(\hat{f}_x + \varepsilon_x - \varphi(\pi)) \\ &\leq f(x) - m(f(x) - \varphi(\pi)) \end{aligned}$$

d'où (4.8). Il reste alors appliquer le Théorème 4.4.3.

□

Remarquons que (5.12) est équivalente à la condition

$$\hat{f} - \varphi(\pi) \leq (1 - m)(\hat{f}_x + \varepsilon_x - \varphi(\pi)) - \varepsilon - \varepsilon_x. \tag{5.13}$$

Donc, à la place d'utiliser (5.10) et le Théorème 5.1.1 pour obtenir (4.10), on peut combiner la condition

$$\varepsilon < (1 - m)(\hat{f}_x + \varepsilon_x - \varphi^j(y^j)) - \varepsilon - \varepsilon_x \quad \forall j \tag{5.14}$$

avec le Théorème 5.1.1 afin de conclure que pour j suffisamment grand

$$f(y^j) - \varphi^j(y^j) \leq (1 - m)(\hat{f}_x + \varepsilon_x - \varphi^j(y^j)) - \varepsilon - \varepsilon_x$$

ce qui implique (5.13) car $\hat{f} \leq f(\pi)$. Nous maintenons (5.14) en diminuant ε à chaque fois que l'expression en question n'est pas vérifiée et nous diminuons si nécessaire ε_x afin de garder le membre de droite de (5.14) strictement positif. Le Théorème 5.2.1 décrit les règles pour choisir ε^j et ε_x^j afin d'obtenir la convergence. La démonstration de ce théorème nécessite un lemme technique

similaire à la deuxième partie du Théorème 4.4.3.

Lemme 5.2.2 *Soient $f : \mathbb{R}^n \to \mathbb{R}$ convexe, $\nu \in (0, \infty)$ et $\kappa > 0$. Supposons que $x \in \mathbb{R}^n$ ne minimise pas f. Alors, il existe $\rho > 0$ tel que pour toutes fonctions convexes $\varphi : \mathbb{R}^n \to \mathbb{R}$ qui minorent f et $\forall M \in S^n$ où $0 < M \leq \nu I$ avec $\pi := p_M^{\varphi}(x)$, l'inégalité $f(\pi) - \varphi(\pi) \leq \rho$ implique*

$$f(\pi) - \varphi(\pi) \leq \kappa \, (f(x) - \varphi(\pi)). \tag{5.15}$$

Preuve : Tout comme dans (4.9), nous avons $f_M(x) \geq \varphi(\pi)$ d'où

$$\gamma := f(x) - f_{\nu I}(x) \leq f(x) - f_M(x) \leq f(x) - \varphi(\pi)$$

et $\gamma > 0$ car x ne minimise pas f. Posons $\rho = \kappa\gamma$ pour établir l'implication désirée.

□

Théorème 5.2.1 *Supposons que les hypothèses du Théorème 4.4.4 où la condition (4.15) est remplacée par (5.11) sont vérifiées. Soient* $\alpha \in (0,1)$, $m \in (0,1)$ *et* $\beta \in (0, \frac{\alpha}{1-\alpha m}]$. *Posons*

$$\sigma := \frac{\alpha(1-m)}{1+\alpha+m\alpha\beta}.$$

Supposons que $\{\varepsilon^j\}$ *et* $\{\varepsilon_x^j\}$ *satisfont*

$$\varepsilon^j = \min\{\varepsilon^{j-1}, \sigma(\hat{f}_x^j - \varphi^j(y^j))\}, \tag{5.16}$$

$$\varepsilon_x^j \in [0, \beta\varepsilon^j], \tag{5.17}$$

où $\varepsilon^0 \geq 0$. *Supposons que les suites de valeurs approximées* $\{\hat{f}^j\}$ *et* $\{\hat{f}_x^j\}$ *satisfont*

$$\hat{f}^j \leq f(y^j) \leq \hat{f}^j + \varepsilon^j, \tag{5.18}$$

$$\hat{f}_x^j \leq f(x) \leq \hat{f}_x^j + \varepsilon_x^j. \tag{5.19}$$

Supposons aussi que les modèles $\{\varphi^j\}$ *satisfont*

$$\varphi^j(x) = \hat{f}_x^j. \tag{5.20}$$

Alors, les assertions suivantes sont vérifiées :

1. *Si* x *ne minimise pas* f *alors (5.12) est vérifiée avec* $\varphi = \varphi^J$, $\pi = y^J$, $\hat{f} = \hat{f}^J$, $\hat{f}_x = \hat{f}_x^J$, $\varepsilon = \varepsilon^J$ *et* $\varepsilon_x = \varepsilon_x^J$ *pour un certain* $J \in I\!N$.

2. *Si* x *minimise* f *alors* $y^j \to x$ *et* $\varepsilon^j, \varepsilon_x^j \to 0$.

Preuve :

1. Par définition de y^j, nous savons que $\varphi^j(y^j) \leq \varphi^j(x)$ d'où (5.20) implique que la suite $\{\varepsilon^j\}$ est positive. Cette suite est décroissante par définition. Par conséquent, la suite $\{\varepsilon^j\}$ s'approche d'une limite $\bar{\varepsilon} \geq 0$.

   Supposons que $\bar{\varepsilon} > 0$. Soit $\theta \in (0,1)$ tel que $\theta > \alpha$. Alors, pour tout j suffisamment grand, $\varepsilon^j \leq \bar{\varepsilon}/\theta$ et les hypothèses du Théorème 5.1.1

sont vérifiées avec $\varepsilon = \bar{\varepsilon}/\theta$. Par conséquent,

$$\limsup_{j\to\infty} (f(y^j) - \varphi^j(y^j)) \leq \bar{\varepsilon}/\theta < \bar{\varepsilon}/\alpha$$

d'où il existe un $J \in \mathbb{N}$ tel que

$$f(y^J) - \varphi^J(y^J) \leq \bar{\varepsilon}/\alpha \leq \varepsilon^J/\alpha. \tag{5.21}$$

Bornons ε^j/α pour tout j de la façon suivante :

$$\begin{aligned}
(5.16) \quad &\Rightarrow \varepsilon^j \leq \sigma(\hat{f}_x^j - \varphi^j(y^j)) \\
&\Leftrightarrow (1 + \alpha + m\alpha\beta)\varepsilon^j \leq \alpha(1-m)(\hat{f}_x^j - \varphi^j(y^j)) \\
&\Rightarrow (1+\alpha)\varepsilon^j + m\alpha\varepsilon_x^j \leq \alpha(1-m)(\hat{f}_x^j - \varphi^j(y^j)) \quad \text{(par (5.17))} \\
&\Leftrightarrow \tfrac{\varepsilon^j}{\alpha} \leq (1-m)(\hat{f}_x + \varepsilon_x^j - \varphi^j(y^j)) - \varepsilon^j - \varepsilon_x^j.
\end{aligned}$$

Par (5.18), nous avons que $\hat{f}^J - \varphi^J(y^J) \leq f(y^J) - \varphi^J(y^J)$. Nous pouvons alors conclure que (5.13) (ou (5.12)) est vérifiée avec les substitutions prescrites dans l'énoncé du théorème.

Supposons que $\bar{\varepsilon} = 0$. Soit $\kappa > 0$ non spécifié temporairement et $\rho > 0$ obtenu grâce au Lemme 5.2.2. Alors $\bar{\varepsilon} < \frac{1}{2}\rho$ et pour tout j suffisamment grand, nous avons que $\varepsilon^j \leq \frac{1}{2}\rho$ d'où les hypothèses du Théorème 5.1.1 sont vérifiées avec $\varepsilon = \frac{1}{2}\rho$. Par conséquent, nous obtenons que

$$\limsup_{j\to\infty} (f(y^j) - \varphi^j(y^j)) \leq \frac{1}{2}\rho < \rho$$

d'où il existe $J \in \mathbb{N}$ tel que $f(y^J) - \varphi^J(y^J) \leq \rho$. Appliquons alors le Lemme 5.2.2 pour obtenir

$$\begin{aligned}
\hat{f}^J - \varphi^J(y^J) &\leq f(y^J) - \varphi^J(y^J) \\
&\leq \kappa(f(x) - \varphi^J(y^J)) \\
&\leq \kappa(\hat{f}_x^J + \varepsilon_x^J - \varphi^J(y^J)). \quad \text{(par (5.19))}
\end{aligned}$$

Afin de compléter la preuve, choisissons κ de façon à ce que l'expression ci-dessus implique (5.13) (ou (5.12)). Nous avons besoin que

$$\begin{aligned}
& \kappa \leq 1 - m - \frac{\varepsilon^J + \varepsilon_x^J}{\hat{f}_x^J + \varepsilon_x^J - \varphi^J(y^J)} && \text{(par (5.17))} \\
\Leftarrow \quad & \kappa \leq 1 - m - \frac{(1+\beta)\varepsilon^J}{\hat{f}_x^J + \varepsilon_x^J - \varphi^J(y^J)} \\
\Leftarrow \quad & \kappa \leq 1 - m - (1+\beta)\frac{\varepsilon^J}{\hat{f}_x^J - \varphi^J(y^J)} && \text{(par (5.16))} \\
\Leftarrow \quad & \kappa \leq 1 - m - (1+\beta)\sigma.
\end{aligned}$$

Si nous prenons $\kappa = 1 - m - (1+\beta)\sigma$ alors $\kappa > 0$ car $\beta \leq \alpha/(1-m\alpha)$ implique

$$\begin{aligned}
(1+\beta)\sigma &= (1+\beta)\frac{\alpha(1-m)}{1+\alpha+m\alpha\beta} \\
&\leq (1+\alpha+m\alpha\beta)\frac{\alpha(1-m)}{1+\alpha+m\alpha\beta} \\
&< 1 - m.
\end{aligned}$$

Par conséquent, avec cette valeur de κ, nous obtenons (5.13) avec les substitutions prescrites dans l'énoncé du théorème.

2. Nous devons avoir $\varepsilon^j \to 0$ car sinon, la première partie de la preuve de (1) s'applique et un pas de descente est généré ce qui est impossible puisque x minimise f. Remarquons que la deuxième partie de la preuve de (1) ne s'applique pas car le Lemme 5.2.2 employé dans cette partie nécessite que x ne minimise pas f. La convergence de $\{\varepsilon_x^j\}$ suit immédiatement de (5.17).
   Soit $\varepsilon > 0$ arbitraire. Pour tout j suffisamment grand, nous avons que $\varepsilon^j \leq \varepsilon$. Les hypothèses du Corollaire 5.1.1 sont donc vérifiées pour ε et on l'applique avec $\sigma = 0$ pour conclure que $\|\bar{y} - x\| \leq \sqrt{\varepsilon/\lambda_{\min}(M)}$ pour tout point d'accumulation $\bar{y}$ de $\{y^j\}$. Puisque ε est arbitraire, nous obtenons $y^j \to x$.

□

Mentionnons que dans une implémentation pratique, on ne poserait pas ε_x^j à $\beta\varepsilon^j$ à chaque itération puisque cela empêcherait un raffinement de l'estimation de $f(x)$ à chaque itération intérieure. A la place, si $\varepsilon_x^{j-1} > \beta\varepsilon^j$ alors on donnerait une valeur à ε_x^j bien plus faible que $\beta\varepsilon^j$, disons 1/10 de cette valeur, afin de réduire la fréquence de réévaluation.
Pour terminer, remarquons que si x ne minimise pas f, alors les tolérances prescrites pas (5.16) et (5.17) ont des bornes inférieures strictement positives.

Proposition 5.2.1 *Supposons que les hypothèses du Théorème 5.2.1 sont vérifiées. Si x ne minimise pas f, alors $\lim_{j\to\infty} \varepsilon^j > 0$. De plus, si $\varepsilon_x^j \geq \gamma\varepsilon^j$ pour un certain $\gamma \in (0, \beta]$, alors $\lim_{j\to\infty} \varepsilon_x^j > 0$.*

Preuve :
La deuxième conclusion suit immédiatement de la première. Prouvons la première conclusion par l'absurde en supposant que $\varepsilon^j \to 0$. La condition (5.16) implique l'existence d'une sous-suite $\{\varepsilon^{j_k}\}$ telle que

$$\begin{aligned} \varepsilon^{j_k} &= \sigma[\hat{f}_x^{j_k} - \varphi^{j_k}(y^{j_k})] \\ &\geq \sigma[f(x) - \varepsilon_x^{j_k} - \varphi^{j_k}(y^{j_k})]. \qquad \text{(par 5.19)} \end{aligned}$$

Par conséquent, par la condition (5.17), nous obtenons que

$$\varepsilon^{j_k} \geq \frac{\sigma}{1+\beta\sigma}[f(x) - \varphi^{j_k}(y^{j_k})].$$

Par ailleurs,

$$\begin{aligned} \varphi^{j_k}(y^{j_k}) &\leq \varphi^{j_k}(y^{j_k}) + \tfrac{1}{2}\|y^{j_k} - x\|_M^2 \\ &= \tilde{\varphi}^{j_k}(y^{j_k}) \\ &\leq \tilde{\varphi}^{j_k}(x) \qquad \text{(Par déf. de } y^{j_k}\text{)} \\ &= \varphi^{j_k}(x) \\ &\leq f(x). \end{aligned}$$

Puisque $\varepsilon^{j_k} \to 0$, les inégalités ci-dessus deviennent des égalités à la limite. En particulier, ceci implique que x est un point d'accumulation de la suite $\{y^j\}$. D'autre part, le Corollaire 5.1.1 s'applique pour tout $\varepsilon > 0$ car $\varepsilon^j \to 0$. On conclut alors que x est minimum de f ce qui contredit les hypothèses.

□

5.3 Algorithme du point proximal inexact

Nous concluons ce chapitre en résumant les modifications apportées à l'algorithme du point proximal approximé. De nouveau, la convergence de l'algorithme est démontrée dans un unique théorème et un théorème supplémentaire nous prouve qu'un point ε-stationnaire peut être obtenu en un temps fini.

Algorithme du Point Proximal Inexact
Soient $\nu_{\max} > 0$, $\varepsilon_{\text{tol}} \geq 0$, $m \in (0,1)$, $\alpha \in (0,1)$, $\beta \in (0, \frac{\alpha}{1-\alpha m}]$ et un point de départ $x^0 \in \mathbb{R}^n$.
1. Poser $\sigma = \frac{\alpha(1-m)}{1+\alpha+m\alpha\beta}$.
2. Pour $k = 0, 1, ...$
3. Choisir $\varepsilon^0 \geq 0$ et ε_x^0 satisfaisant (5.17).
4. Calculer $\hat{f}_x^0 \in [f(x^k) - \varepsilon_x^0, f(x^k)]$.
5. Choisir un modèle φ^1 minorant f tel que $\varphi^1(x^k) = \hat{f}_x^0$.
6. Choisir M_k tel que $0 < M_k \leq \nu_{\max} I$.
7. Poser $j = 0$.
8. Répéter
9. $j \leftarrow j + 1$.
10. Calculer $y^j = p_{M_k}^{\varphi^j}(x^k)$ et $g^j = s_{M_k}^{\varphi^j}(x^k)$.
11. Choisir ε^j et ε_x^j satisfaisant (5.16)-(5.17) et calculer $\hat{f}^j$ et $\hat{f}_x^j$ satisfaisant

$$\hat{f}^j \in [f(y^j) - \varepsilon^j, f(y^j)],$$
$$\hat{f}_x^j \in [f(x^k) - \varepsilon_x^j, f(x^k)].$$

12. Si $\hat{f}^j + \varepsilon^j - \varphi^j(y^j) \leq \varepsilon_{\text{tol}}$ et $\|g^j\| \leq \varepsilon_{\text{tol}}$ alors
13. Poser $\bar{x} = y^j$ et STOP $\Rightarrow$ $\bar{x}$ est un point ε-stationnaire.
14. Fin si
15. Choisir un modèle φ^{j+1} satisfaisant (4.14),(4.16) et tel que

$$\varphi^{j+1}(y) \geq \hat{f}^j + < \hat{s}(y^j), y - y^j > \quad \forall y \in \mathbb{R}^n$$

où $\hat{s}(y^j) \in \partial_{\varepsilon^j} f(y^j)$.

16. Jusqu'à ce que $\hat{f}^j + \varepsilon^j \leq \hat{f}_x^j - m(\hat{f}_x^j + \varepsilon_x^j - \varphi^j(y^j))$
17. Poser $x^{k+1} = y^j$.
18. Fin pour

Théorème 5.3.1 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ convexe et soit la suite $\{x^k\}$ donnée par l'algorithme du point proximal inexact avec $\varepsilon_{tol} = 0$. Si la suite $\{x^k\}$ est finie et si x^K est le dernier élément de la suite, x^K minimise f. Sinon, si $\{x^k\}$ est bornée, ses points d'accumulation minimisent f et $f(x^k) \to \min_x f(x)$.*

Preuve :
Pour commencer, mentionnons que la troisième contrainte à la ligne 15 implique (5.11) et que la condition d'arrêt à la ligne 16 est exactement (5.12).

1. La première conclusion suit simplement de la contraposée du Théorème 5.2.1 (1) puisque la condition (5.12) n'est jamais satisfaite pour $x = x^K$.
2. Supposons que $\{x^k\}$ est infinie. Elle satisfait alors la condition de la ligne 16 pour tout k d'où, par le Lemme 5.2.1, la suite satisfait (4.4). Par ailleurs, (4.7) est vérifiée car $M_k \leq \nu_{\max} I$ pour tout k. Il reste à appliquer le Théorème 4.4.2 pour compléter la preuve.

□

Théorème 5.3.2 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ convexe et bornée inférieurement. Soit $\varepsilon_{tol} > 0$. Alors, l'algorithme du point proximal inexact se termine en un temps fini avec un point ε_{tol}-stationnaire $\bar{x}$.*

Preuve :
La preuve de ce théorème est semblable à celle du Théorème 4.4.6 c'est pourquoi nous ne rentrerons pas dans les détails. Remarquons que

$$g^j \in \partial_{\eta_j} f(y^j) \text{ où } \eta_j := f(y^j) - \varphi^j(y^j) \leq \hat{f}^j + \varepsilon^j - \varphi^j(y^j). \tag{5.22}$$

Par conséquent, si l'algorithme se termine (à la ligne 12), l'expression (5.22) nous indique que $\bar{x} = y^j$ est nécessairement un point ε_{tol}-stationnaire. Il ne nous reste donc plus qu'à prouver la terminaison de l'algorithme :

Procédons par contradiction, supposons que l'algorithme ne se termine pas. Dans ce contexte, il y a deux cas à considérer : $\{x^k\}$ est finie ou $\{x^k\}$ est infinie. Dans les deux cas, il faut montrer que la condition d'arrêt à la ligne 12 est finalement satisfaite en un certain point afin de contredire l'hypothèse

de non terminaison. Nous ne considérons que le cas où $\{x^k\}$ est finie.
Supposons donc que $\{x^k\}$ est finie et considérons la suite infinie $\{y^j\}$ (car l'algorithme ne se termine pas) de l'itération extérieure finale K. Par la contraposée du Théorème 5.2.1 (1), nous avons que x^K minimise f d'où, par le Théorème 5.2.1 (2), $y^j \to x^K$ et $\epsilon^j \to 0$. Il reste à appliquer le Théorème 5.1.1 pour tout $\epsilon > 0$ pour conclure que $\eta_j \to 0$. D'autre part, $y^j \to x^K$ implique $g^j \to 0$. Par conséquent, la condition de la ligne 12 doit être en fin de compte satisfaite.

□

5.4 Application : LMI à grande échelle

5.4.1 Définitions de LMI et SDP

Considérons une fonction affine $F : \mathbb{R}^n \to S^m$. Une telle fonction peut être écrite de la façon suivante :

$$F(x) = F_0 + x_1 F_1 + ... + x_n F_n \tag{5.23}$$

où $F_i \in S^m$, $i = 0, ..., n$. L'affirmation $F(x) \leq 0$ est un type d'inégalité matricielle linéaire (LMI) et exprime une contrainte convexe.
Un problème d'optimisation avec une fonction coût linéaire et une contrainte de type LMI

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & c^T x \\ \text{s.c.} & F(x) \leq 0 \end{array} \tag{5.24}$$

est un programme semidéfini (SDP) où $c \in \mathbb{R}^n$ est le vecteur coût. Comme classe de programmes non linéaires, les programmes semidéfinis ont deux avantages. Premièrement, les exemples pratiques de SDP sont abondants. Ils incluent un grand nombre de programmes convexes non linéaires et tous les problèmes de programmation linéaire. En effet, si les F_i dans (5.23) sont des matrices diagonales alors F est une matrice diagonale dans (5.24). Dès lors, $F(x) \leq 0$ correspond à m contraintes linéaires d'inégalité. Deuxièmement, les SDPs peuvent être résolus en un temps polynomial. Un algorithme en temps polynomial (algorithme qui résout le problème en un nombre d'étapes borné par une fonction polynomiale de la taille du problème) est considéré comme efficace théoriquement bien que l'efficacité pratique dépend aussi de l'ordre du polynôme.
Dans sa thèse, A. Miller présente plusieurs exemples de SDP ayant deux caractéristiques en commun. Premièrement, les exemples pratiques de ces problèmes peuvent être tellement grands que les méthodes de point intérieur

deviennent trop coûteuses en dépit de la borne polynomiale sur le temps d'exécution. Deuxièmement, chaque problème possède une structure particulière pouvant être exploitée pour accélérer la résolution.

5.4.2 Méthodes au valeur propre

Dans sa thèse, A. Miller s'intéresse uniquement à la résolution du problème d'admissibilité $F(x) \leq 0$ et se base sur les méthodes au valeur propre. Cette approche consiste à minimiser

$$\bar{\lambda}(x) := \lambda_1(F(x)).$$

Si $\min_x \bar{\lambda}(x) \leq 0$ alors le minimum fournit un point admissible sinon le problème est non admissible. Il est évident que l'on peut stopper le processus de minimisation dès que $\bar{\lambda}(x) \leq 0$.
L'avantage des méthodes au valeur propre est qu'une représentation implicite de F peut être utilisée dans un algorithme itératif (algorithme de Lanczos) pour estimer $\bar{\lambda}(x)$. Cela signifie que la matrice $F(x)$ n'est jamais véritablement construite, la matrice $F(x)$ est présente à travers des produits de matrice-vecteur $F(x)q$ et à travers des solutions de systèmes linéaires déplacés $[\sigma I - F(x)]y = b$. Notons que ces opérations sont rapidemment réalisées pour les problèmes structurés décrits dans la thèse de A. Miller.
L'inconvénient de cette approche est qu'elle nécessite des techniques spéciales d'optimisation. En effet, la fonction $\bar{\lambda}(.)$ est convexe et non nécessairement différentiable. Par ailleurs, la fonction $\bar{\lambda}(.)$ et ses sous-gradients ne peuvent être évalués exactement mais ils peuvent être estimés avec n'importe quel degré de précision à l'aide d'une méthode itérative. La méthode du point proximal inexact est donc adaptée à ce type d'approche. Notons que le chapitre 5 de [6] décrit une adaptation de la méthode du point proximal inexact à la minimisation de $\bar{\lambda}(x) := \lambda_1(F(x))$. Il décrit la fonction modèle, comment mettre à jour le modèle, comment mettre à jour la métrique...

Chapitre 6

Méthode Faisceau Inexacte et Analyse de Stabilité

Nous considérons toujours le problème

$$\min_x \{f(x) \mid x \in \mathbb{R}^n\} \tag{6.1}$$

où $f : \mathbb{R}^n \to \mathbb{R}$ est convexe et non nécessairement différentiable. Nous nous intéressons toujours au cas où étant donné un point, nous ne disposons que d'une valeur approximée de f et d'une valeur approximée de l'un de ses sous-gradients. Plus précisément, étant donnés $x \in \mathbb{R}^n$ et $\tilde{\varepsilon} > 0$, nous supposons que l'on peut trouver $\tilde{f} \in \mathbb{R}$ et $y \in \mathbb{R}^n$ tels que

$$f(x) \geq \tilde{f} \geq f(x) - \tilde{\varepsilon},$$

$$f(z) \geq \tilde{f} + < y, z - x > \quad \forall z \in \mathbb{R}^n.$$

Ce chapitre consiste à présenter une méthode de type faisceau basée sur ses hypothèses. Malgré le fait que nous nous situons dans un contexte comparable à celui de la méthode du point proximal inexact, la méthode faisceau qui sera développée est d'un grand intérêt. En effet, elle permet de répondre aux questions suivantes :

. Etant donné un niveau d'optimalité $\Delta_{\text{opt}} > 0$ désiré pour le problème (6.1), avec quelle précision doivent être évalués f et ses sous-gradients afin de garantir une terminaison par un point satisfaisant la tolérance donnée Δ_{opt} ?

. Etant donnée une erreur d'approximation non nulle qui ne tend pas nécessairement vers zéro, quelles sont les propriétés de convergence ? Quel type de solutions approximées au problème (6.1) peut-on obtenir et comment dépendent-elles de l'erreur d'approximation ?

Questions auxquelles la méthode du point proximal inexact était incapable de répondre.

6.1 Algorithme de Solodov

L'algorithme considéré par la suite est, excepté le fait d'être basé sur des données inexactes et d'utiliser une condition d'arrêt spéciale, une méthode faisceau standard où la taille du modèle est contrôlée par la technique d'agrégation. Néanmoins, il est important de détailler nos notations.
L'ensemble des approximations linéaires habituelles de la fonction sont collectées dans l'ensemble B_k^c et sont notées $l_i(x)$, $i \leq k$. L'ensemble des approximations linéaires obtenues après agrégation sont collectées dans l'ensemble B_k^a et sont notées $l_i^a(x)$, $i \leq k$. Les deux ensembles I_k^c et I_k^a contiennent les indices d'itération des membres de B_k^c et de B_k^a respectivement.
Quand le nombre d'éléments dans $B_k^c \cup B_k^a$ atteint la borne supérieure $B_{\max}$, deux éléments ou plus sont supprimés du faisceau et sont remplacés par la pièce d'agrégation (étape 8 de l'algorithme). Ceci contrôle la complexité de l'approximation de f par les plans sécants i.e. φ_k donnée par (6.6). De cette manière, le sous-problème (6.5) à l'étape 2 de l'algorithme reste traitable efficacement. Le sous-problème (6.5) est résolu via une méthode de programmation quadratique appliquée à son dual (voir (6.12) et le Lemme 6.2.1) dont la dimension est précisément $|B_k^c| + |B_k^a|$.
Nous ne donnons pas de règle particulière concernant le choix du paramètre γ_k à l'étape 2 de l'algorithme. Rappelons tout de même que ce choix est important pour l'efficacité pratique des méthodes faisceaux.
A l'étape 3 de l'algorithme, la valeur de la "diminution prédite" est calculée et la solution de (6.5) est acceptée comme étant le meilleur itéré si la véritable diminution est supérieure au produit de $\sigma \in (0,1)$ et de la "diminution prédite" (étape 6 de l'algorithme). Les indices de tels "pas de diminution" sont collectés dans l'ensemble K_d et la meilleure valeur (approximée) de f est notée $\bar{f}_k$.
Par la Proposition 6.2.1, si le critère d'arrêt (6.8) est vérifié à l'itération k, nous obtenons que

$$d^k \in \partial_{\varepsilon_k} f(x^k) \tag{6.2}$$

tel que

$$\Delta_k := \frac{1}{2\gamma_k}\|d^k\|^2 + \varepsilon_k \leq \Delta_{\text{opt}}, \tag{6.3}$$

où d^k est calculé en utilisant la solution de (6.12) (dual de (6.5), voir Lemme 6.2.1).
Dans le Théorème 6.2.1, nous montrons que notre critère d'arrêt sera finalement satisfait si les approximations sont contrôlées par la règle suivante :

$$\frac{1-\sigma}{2(2-\sigma)}\Delta_{\text{opt}} > \limsup_k \left(\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\}\right) \tag{6.4}$$

où σ est le paramètre utilisé dans le test de diminution (étape 6 de l'algorithme). Autrement dit, si les approximations de la fonction objectif et de ses sous-gradients sont suffisamment précises au sens de (6.4), alors l'algorithme se termine après un nombre fini d'itérations en un point x^k vérifiant (6.2) et (6.3). Notons que la précision requise est finie ($\tilde{\varepsilon}_k$ ne doit pas tendre vers 0) et que l'expression (6.4) indique clairement le degré de précision nécessaire pour obtenir la solution approximée désirée de (6.1). Il est maintenant temps de présenter l'algorithme de Solodov.

Algorithme de Solodov
Choisir $\sigma \in (0,1)$, $x^0 \in \mathbb{R}^n$ et $B_{\max} \geq 2$.
Poser $K_d, B_{-1}^c, B_0^a, I_0^a = \emptyset$.
Calculer $\tilde{f}_0 \in \mathbb{R}$ et $y^0 \in \mathbb{R}^n$ tels que $f(x^0) \geq \tilde{f}_0 \geq f(x^0) - \tilde{\varepsilon}_0$, $\tilde{\varepsilon}_0 \geq 0$, $f(x) \geq l_0(x) := \tilde{f}_0 + < y^0, x - x^0 > \forall x \in \mathbb{R}^n$.
Poser $z^0 := x^0$, $\tilde{f}_0 := \tilde{f}_0$ et $k := 0$.

1. Ajouter la nouvelle pièce à l'approximation de f.
   Poser

$$B_k^c := B_{k-1}^c \cup \{l_k(x)\}, \quad l_k(x) := \tilde{f}_k + < y^k, x - z^k >,$$

$$I_k^c := \{0 \leq i \leq k \mid l_i(x) \in B_k^c\}.$$

2. Minimiser l'approximation de f.
   Choisir $\gamma_k > 0$ et calculer z^{k+1} solution de

$$\min_x \{\varphi_k(x) + \frac{\gamma_k}{2}\|x - x^k\|^2 \mid x \in \mathbb{R}^n\}, \tag{6.5}$$

   où

$$\varphi_k(x) := \max\{\max\{l_i(x) \mid i \in I_k^c\},\ \max\{l_i^a(x) \mid i \in I_k^a\}\}. \tag{6.6}$$

3. Calculer la valeur de la "diminution prédite".

$$\delta_k := \bar{f}_k - \varphi_k(z^{k+1}) - \frac{\gamma_k}{2}\|z^{k+1} - x^k\|^2. \tag{6.7}$$

4. Condition d'arrêt.
   Stop si

$$\delta_k + 2\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\} \leq \Delta_{opt}. \tag{6.8}$$

   Sinon, aller à l'étape 5.

5. Approximer les valeurs de f et de ses sous-gradients en z^{k+1}.
   Calculer $\tilde{f}_{k+1} \in \mathbb{R}$ et $y^{k+1} \in \mathbb{R}^n$ tel que

$$\begin{array}{c} f(z^{k+1}) \geq \tilde{f}_{k+1} \geq f(z^{k+1}) - \tilde{\varepsilon}_{k+1},\ \tilde{\varepsilon}_{k+1} \geq 0, \\ f(x) \geq l_{k+1}(x) := \tilde{f}_{k+1} + < y^{k+1}, x - z^{k+1} > \ \forall x \in \mathbb{R}^n. \end{array} \tag{6.9}$$

6. Test de diminution.
   Si

$$\bar{f}_k - \tilde{f}_{k+1} - \tilde{\varepsilon}_{k+1} \geq \sigma\delta_k, \tag{6.10}$$

   poser $x^{k+1} := z^{k+1}$, $\bar{f}_{k+1} := \tilde{f}_{k+1}$, $K_d := K_d \cup \{k+1\}$ et aller à l'étape 8.

7. Pas nul.
   Poser $x^{k+1} := x^k$ et $\bar{f}_{k+1} := \bar{f}_k$.

8. Gestion de la complexité de l'approximation de f.
   Si $|B_k^c| + |B_k^a| < B_{\max}$, alors poser $B_{k+1}^a := B_k^a$ et aller à l'étape 9.
   Sinon, choisir $C_k \subset B_k^c \cup B_k^a$ tel que $|C_k| \geq 2$, $l_{k_0}(x) \notin C_k$, où $k_0 = \max\{i \mid i \in K_d\}$. Poser

$$\begin{array}{c} B_k^c := B_k^c \setminus C_k,\ B_{k+1}^a := (B_k^a \setminus C_k) \cup \{l_k^a(x)\}, \\ l_k^a(x) := \varphi_k(z^{k+1}) + \gamma_k\|z^{k+1} - x^k\|^2 + \gamma_k < x^k - z^{k+1}, x - x^k >, \\ I_{k+1}^a := \{0 \leq i \leq k+1 \mid l_i^a(x) \in B_{k+1}^a\}. \end{array}$$

9. Poser $k := k+1$ et aller à l'étape 1.

6.2 Propriétés de convergence

Nous commençons par étudier la caractérisation duale de la solution de (6.5). Soient $(\alpha_i, v^i, u^i) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$, $i = 1, ..., |B_k^c| + |B_k^a|$, les données qui définissent φ_k dans (6.6). Nous utiliserons la représentation

$$\begin{array}{c} \alpha_i + < v^i, x - u^i > = l_i(x) \in B_k^c,\ i \in I_k^c, \\ \alpha_i + < v^i, x - u^i > = l_i^a(x) \in B_k^a,\ i \in I_k^a. \end{array} \tag{6.11}$$

Formellement parlant, certains indices se répètent dans (6.11) puisque $I_k^c \cap I_k^a \neq \emptyset$. Malgré cela, nous utiliserons cette représentation puisqu'elle ne mène à aucune réelle confusion et qu'elle simplifie singulièrement les notations. Soit I_k l'union des indices qui définissent φ_k ($|I_k| = |B_k^c| + |B_k^a|$) et considérons le programme quadratique suivant

$$\begin{array}{ll} \max_\lambda & -\frac{1}{2\gamma_k}\|\sum_{i\in I_k}\lambda_i v^i\|^2 + \sum_{i\in I_k}\lambda_i(\alpha_i + < v^i, x^k - u^i >) \\ \text{s.c.} & \lambda \geq 0 \\ & \sum_{i\in I_k}\lambda_i = 1 \end{array} \tag{6.12}$$

où $\lambda \in \mathbb{R}^{|I_k|}$.

Lemme 6.2.1 *Soit $\bar{\lambda}^k$ une solution de (6.12) et notons*

$$d^k := \sum_{i\in I_k} \bar{\lambda}_i^k v^i.$$

Alors, les assertions suivantes sont vérifiées :

$$z^{k+1} = x^k - \frac{1}{\gamma_k} d^k, \tag{6.13}$$

$$d^k \in \partial\varphi_k(z^{k+1}), \tag{6.14}$$

$$d^k \in \partial_{\varepsilon_k} f(x^k), \tag{6.15}$$

où $\varepsilon_k = \varepsilon_k^c + \varepsilon_k^a \geq 0$,

$$\varepsilon_k^c = \sum_{i\in I_k^c} \bar{\lambda}_i^k (f(x^k) - f(z^i) - < y^i, x^k - z^i > + \tilde{\varepsilon}_i) \geq 0,$$

$$\varepsilon_k^a = \sum_{i\in I_k^a} \bar{\lambda}_i^k (f(x^k) - \varphi_i(z^{i+1}) - \frac{1}{\gamma_i}\|d^i\|^2 - < d^i, x^k - x^i >) \geq 0.$$

Preuve :
Le problème (6.5) est équivalent au programme quadratique convexe

$$\min_{x,t} \{t + \frac{\gamma_k}{2}\|x - x^k\|^2 \mid l_i(x) \leq t,\ i \in I_k^c,\ l_i^a(x) \leq t,\ i \in I_k^a\}. \tag{6.16}$$

La fonction lagrangienne associée au problème (6.16) est

$$L(x,t,\lambda) = t + \frac{\gamma_k}{2}\|x - x^k\|^2 + \sum_{i\in I_k} \lambda_i(\alpha_i + < v^i, x - u^i > -t).$$

La fonction duale est $d(\lambda) = \min_{x,t} L(x,t,\lambda)$. Pour trouver le minimum de $L(x,t,\lambda)$, nous résolvons le système

$$\nabla_t L(x,t,\lambda) = 1 - \textstyle\sum_{i\in I_k} \lambda_i = 0,$$

$$\nabla_x L(x,t,\lambda) = \gamma_k(x - x^k) + \textstyle\sum_{i\in I_k} \lambda_i v^i = 0.$$

Le problème dual de (6.5) est donc

$$\begin{array}{ll} \max_{x,t,\lambda} & t + \frac{\gamma_k}{2}\|x - x^k\|^2 + \sum_{i\in I_k} \lambda_i(\alpha_i + < v^i, x - u^i > -t) \\ \text{s.c.} & x = x^k - \frac{1}{\gamma_k}\sum_{i\in I_k} \lambda_i v^i \\ & \sum_{i\in I_k} \lambda_i = 1 \\ & \lambda \geq 0, \end{array} \tag{6.17}$$

et devient le problème (6.12) après élimination des variables x et t.
L'assertion (6.13) du Lemme est maintenant évidente par les contraintes de (6.17), l'unicité de z^{k+1} et la forte dualité.
Puisque z^{k+1} est solution optimale de (6.5), nous avons que

$$\begin{aligned} 0 \quad &\in \partial\varphi_k(z^{k+1}) + \tfrac{\gamma_k}{2}\nabla\|x - x^k\|^2(z^{k+1}) \\ &= \partial\varphi_k(z^{k+1}) + \gamma_k(z^{k+1} - x^k) \\ &= \partial\varphi_k(z^{k+1}) - d^k. \end{aligned}$$

L'assertion (6.14) est donc vérifiée.
En ce qui concerne la dernière assertion, commençons par démontrer par induction que

$$f(x) \geq \varphi_k(x) \quad \forall x \in \mathbb{R}^n,\ \forall k. \tag{6.18}$$

Pour $k = 0$, c'est évident puisque $\varphi_0(x) = l_0(x)$. Supposons que (6.18) est vérifiée pour l'indice k. Si $|B_k^c| + |B_k^a| < B_{\max}$, alors

$$\varphi_{k+1}(x) = \max\{\varphi_k(x),\ l_{k+1}(x)\} \leq f(x),$$

par (6.9) et (6.18). Si $|B_k^c| + |B_k^a| = B_{\max}$, alors

$$\varphi_{k+1}(x) \leq \max\{\varphi_k(x),\ l_k^a(x),\ l_{k+1}(x)\} \leq f(x),$$

car

$$\begin{aligned} l_k^a(x) &= \varphi_k(z^{k+1}) + \gamma_k \|z^{k+1} - x^k\|^2 + \gamma_k < x^k - z^{k+1}, x - x^k > \\ &= \varphi_k(z^{k+1}) + \gamma_k < z^{k+1} - x^k, z^{k+1} - x^k - x + x^k > \\ &= \varphi_k(z^{k+1}) + < d^k, x - z^{k+1} > \leq \varphi_k(x), \end{aligned}$$

par (6.14). Par conséquent, (6.18) est vérifiée.
Les valeurs optimales de (6.5) et (6.12) sont égales par la dualité forte, nous obtenons donc

$$\varphi_k(z^{k+1}) = -\frac{1}{\gamma_k}\|d^k\|^2 + \sum_{i \in I_k} \bar{\lambda}_i^k(\alpha_i + < v^i, x^k - u^i >). \tag{6.19}$$

Par (6.18), nous obtenons pour tout $x \in \mathbb{R}^n$,

$$\begin{aligned} f(x) &\geq \varphi_k(x) \geq \varphi_k(z^{k+1}) + < d^k, x - z^{k+1} > \\ &= -\tfrac{1}{\gamma_k}\|d^k\|^2 + < d^k, x - x^k + \tfrac{1}{\gamma_k} d^k > \\ &\quad + \textstyle\sum_{i \in I_k} \bar{\lambda}_i^k(\alpha_i + < v^i, x^k - u^i >) \\ &= f(x^k) + < d^k, x - x^k > \\ &\quad - \textstyle\sum_{i \in I_k} \bar{\lambda}_i^k(f(x^k) - \alpha_i - < v^i, x^k - u^i >) \\ &\geq f(x^k) + < d^k, x - x^k > \\ &\quad - \textstyle\sum_{i \in I_k^c} \bar{\lambda}_i^k(f(x^k) - f(z^i) - < y^i, x^k - z^i > + \tilde{\varepsilon}_i) \\ &\quad - \textstyle\sum_{i \in I_k^a} \bar{\lambda}_i^k(f(x^k) - \varphi_i(z^{i+1}) - \tfrac{1}{\gamma_i}\|d^i\|^2 - < d^i, x^k - x^i >) \\ &= f(x^k) + < d^k, x - x^k > - \varepsilon_k^c - \varepsilon_k^a, \end{aligned}$$

où la seconde inégalité est une conséquence de (6.14), la première égalité vient de (6.19) et (6.13), et la troisième inégalité vient de (6.11) et (6.9). La quantité ε_k^c est positive car $y^i \in \partial_{\tilde{\varepsilon}_i} f(z^i)$ par (6.9). Il reste à estimer ε_k^a. Par (6.13),(6.14) et (6.18), nous obtenons

$$
\begin{aligned}
\varepsilon_k^a &= \textstyle\sum_{i\in I_k^a} \bar{\lambda}_i^k (f(x^k) - \varphi_i(z^{i+1}) - \frac{1}{\gamma_i}\|d^i\|^2 - < d^i, x^k - x^i >) \\
&= \textstyle\sum_{i\in I_k^a} \bar{\lambda}_i^k (f(x^k) - \varphi_i(z^{i+1}) - \frac{1}{\gamma_i} < d^i, d^i + \gamma_i(x^k - x^i) > \\
&= \textstyle\sum_{i\in I_k^a} \bar{\lambda}_i^k (f(x^k) - \varphi_i(z^{i+1}) - < d^i, x^k - z^{i+1} >) \\
&\geq \textstyle\sum_{i\in I_k^a} \bar{\lambda}_i^k (f(x^k) - \varphi_i(x^k)) \\
&\geq 0.
\end{aligned}
$$

□

Le résultat suivant indique que le critère d'arrêt de l'algorithme garantit la précision désirée dans la solution de (6.1).

Proposition 6.2.1 *Supposons que le critère d'arrêt (6.8) est satisfait à l'itération k. Alors, x^k est une solution approximée de (6.1) au sens de (6.2) et (6.3) où d^k est défini dans le Lemme 6.2.1.*

Preuve :
Soit $k_0 = \max\{i \mid i \in K_d\}$ l'indice du dernier pas de descente avant que la condition (6.8) ne soit satisfaite. Par (6.7) et puisque $\bar{f}_k = \bar{f}_{k_0}$, nous avons que

$$
\begin{aligned}
\delta_k &= \bar{f}_k - \varphi_k(z^{k+1}) - \tfrac{1}{2\gamma_k}\|d^k\|^2 \\
&= \bar{f}_{k_0} - \varphi_k(z^{k+1}) - \tfrac{1}{2\gamma_k}\|d^k\|^2 \\
&= \bar{f}_{k_0} + \tfrac{1}{2\gamma_k}\|d^k\|^2 - \textstyle\sum_{i\in I_k} \bar{\lambda}_i^k (\alpha_i + < v^i, x^k - u^i >) \\
&\geq \tfrac{1}{2\gamma_k}\|d^k\|^2 + \textstyle\sum_{i\in I_k} \bar{\lambda}_i^k (f(x^{k_0}) - \alpha_i - < v^i, x^k - u^i >) - \tilde{\varepsilon}_{k_0}
\end{aligned}
$$

où la troisième égalité et l'inégalité sont des conséquences de (6.19) et (6.9)

respectivement. Puisque $x^{k_0} = x^k$, nous obtenons

$$\begin{aligned}\textstyle\sum_{i\in I_k^c} \bar{\lambda}_i^k (f(x^{k_0}) - \alpha_i - < v^i, x^k - u^i >) &= \textstyle\sum_{i\in I_k^c} \bar{\lambda}_i^k (f(x^k) - \tilde{f}_i \\ &\quad - < y^i, x^k - z^i >) \\ &\geq \textstyle\sum_{i\in I_k^c} \bar{\lambda}_i^k (f(x^k) - f(z_i) \\ &\quad - < y^i, x^k - z^i >) \\ &= \varepsilon_k^c - \textstyle\sum_{i\in I_k^c} \bar{\lambda}_i^k \tilde{\varepsilon}_i \\ &\geq \varepsilon_k^c - \max\{\tilde{\varepsilon}_i \mid i \in I_k^c\}\end{aligned}$$

où la première inégalité vient de (6.9). Par ailleurs, nous savons que

$$\sum_{i\in I_k^a} \bar{\lambda}_i^k (f(x^k) - \alpha_i - < v^i, x^k - u^i >) = \varepsilon_k^a.$$

Nous pouvons alors conclure que

$$\begin{aligned}\delta_k &\geq \tfrac{1}{2\gamma_k}\|d^k\|^2 + \varepsilon_k^c + \varepsilon_k^a - \tilde{\varepsilon}_{k_0} - \max\{\tilde{\varepsilon}_i \mid i \in I_k^c\} \\ &\geq \tfrac{1}{2\gamma_k}\|d^k\|^2 + \varepsilon_k - 2\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\}\end{aligned}$$

où la dernière inégalité suit du fait que $k_0 \in I_k^c$, par l'étape 8 de l'algorithme. La dernière relation implique (6.3) quand (6.8) est vérifiée.

□

Etant donnée la Proposition 6.2.1, il reste à prouver que le critère d'arrêt (6.8) est nécessairement vérifié si la précision des approximations satisfait la condition (6.4).

Théorème 6.2.1 *Supposons que la valeur optimale de (6.1) est finie et que $\gamma_{k+1} \geq \gamma_k$ et $\tilde{\varepsilon}_{k+1} \leq \tilde{\varepsilon}_k$ pour tout k suffisamment grand, $k \notin K_d$. Alors, les assertions suivantes sont mutuellement exclusives :*

1. *L'algorithme exécute un nombre infini d'itérations.*
2. *La condition (6.4) est vérifiée.*

Preuve :
Supposons que la suite $\{x^k\}$ est infinie ce qui signifie que le critère d'arrêt (6.8) n'est jamais satisfait i.e.

$$\delta_k + 2\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\} > \Delta_{\text{opt}} \quad \forall k \tag{6.20}$$

et supposons que la condition (6.4) est vérifiée afin d'obtenir une contradiction.

Pour commencer, supposons qu'il y a un nombre infini de pas de diminution. Pour tout $k+1 \in K_d$, (6.10) et (6.9) implique que

$$\begin{aligned} \sigma\delta_k &\leq \bar{f}_k - (\tilde{f}_{k+1} + \tilde{\varepsilon}_{k+1}) \\ &\leq f(x^k) - f(x^{k+1}). \end{aligned}$$

D'où,

$$\begin{aligned} \sigma \textstyle\sum_{k+1\in K_d} \delta_k &\leq \textstyle\sum_{k+1\in K_d} (f(x^k) - f(x^{k+1})) \\ &\leq f(x^0) - \min\{f(x) \mid x \in R^n\} < +\infty. \end{aligned}$$

En particulier, il suit que $\liminf_{k+1\in K_d} \delta_k \leq 0$. Par (6.20), nous obtenons donc que

$$\limsup_{k+1\in K_d} (\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\}) \geq \frac{\Delta_{\text{opt}}}{2}.$$

Par conséquent, nous obtenons que

$$\begin{aligned} \limsup_k (\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\}) &\geq \limsup_{k+1\in K_d} (\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\}) \\ &\geq \tfrac{\Delta_{\text{opt}}}{2} \\ &\geq \tfrac{(1-\sigma)}{(2-\sigma)} \tfrac{\Delta_{\text{opt}}}{2} \end{aligned}$$

ce qui contredit (6.4).

Pour terminer, supposons que le nombre de pas de diminution est fini. Soit $k_0 = \max\{i \mid i \in K_d\}$ l'indice du dernier pas de diminution tel que $x^k = x^{k_0}$, $\bar{f}_k = \tilde{f}_{k_0}$ pour tout $k \geq k_0$. Pour tout k suffisamment grand, disons $k \geq k_1$,

nous obtenons que

$$\begin{aligned}\bar{f}_k - \delta_k + \tfrac{\gamma_k}{2}\|z^{k+2} - z^{k+1}\|^2 &= \tfrac{\gamma_k}{2}(\|z^{k+1} - x^{k_0}\|^2 + \|z^{k+2} - z^{k+1}\|^2) \\ &\quad +\varphi_k(z^{k+1}) \\ &= \gamma_k < x^{k_0} - z^{k+1}, z^{k+2} - z^{k+1} > \\ &\quad +\tfrac{\gamma_k}{2}\|z^{k+2} - x^{k_0}\|^2 + \varphi_k(z^{k+1}) \\ &\leq < d^k, z^{k+2} - z^{k+1} > +\tfrac{\gamma_{k+1}}{2}\|z^{k+2} - x^{k_0}\|^2 \\ &\quad +\varphi_k(z^{k+1}),\end{aligned}$$

où la première égalité suit de (6.7) et la seule inégalité vient de (6.13) et $\gamma_k \leq \gamma_{k+1}$.
Par ailleurs, si $|B_k^c| + |B_k^a| < B_{\max}$, nous avons que

$$\begin{aligned}\varphi_k(z^{k+1}) + < d^k, z^{k+2} - z^{k+1} > &\leq \varphi_k(z^{k+2}) \\ &\leq \max\{\varphi_k(z^{k+2}), l_{k+1}(z^{k+2})\} \\ &= \varphi_{k+1}(z^{k+2}),\end{aligned}$$

où la première inégalité suit de (6.14). Si $|B_k^c| + |B_k^a| = B_{\max}$, nous avons que

$$\varphi_k(z^{k+1}) + < d^k, z^{k+2} - z^{k+1} > = l_k^a(z^{k+2}) \leq \varphi_{k+1}(z^{k+2}).$$

En combinant les deux cas, nous obtenons que

$$\begin{aligned}\bar{f}_k - \delta_k + \tfrac{\gamma_k}{2}\|z^{k+2} - z^{k+1}\|^2 &\leq \varphi_{k+1}(z^{k+2}) + \tfrac{\gamma_{k+1}}{2}\|z^{k+2} - x^{k_0}\|^2 \\ &= \bar{f}_k - \delta_{k+1},\end{aligned}$$

où l'égalité suit de (6.7). D'où, pour $k \geq k_1$,

$$\delta_k \geq \delta_{k+1} + \frac{\gamma_k}{2}\|z^{k+2} - z^{k+1}\|^2 \geq \delta_{k+1}. \tag{6.21}$$

En particulier, la suite $\{\delta_k\}$ est décroissante (pour $k \geq k_1$) et bornée inférieurement par (6.20) et (6.4). Par conséquent, elle converge :

$$\bar{\delta} = \lim_{k\to\infty} \delta_k. \tag{6.22}$$

Montrons que la suite $\{z^k\}$ est bornée. Nous avons que

$$\begin{aligned}\bar{f}_k - \delta_k + \tfrac{\gamma_k}{2}\|z^{k+1} - x^{k_0}\|^2 &= \varphi_k(z^{k+1}) + \gamma_k\|z^{k+1} - x^{k_0}\|^2 \\ &= \varphi_k(z^{k+1}) + < d^k, x^{k_0} - z^{k+1} > \\ &\leq \varphi_k(x^{k_0}),\end{aligned}$$

où l'inégalité vient de (6.14). Par conséquent, pour k suffisamment grand ($k \geq k_1$), il suit que

$$\begin{aligned}\|z^{k+1} - x^{k_0}\|^2 &\leq \tfrac{2}{\gamma_k}(\delta_k + \varphi_k(x^{k_0}) - \bar{f}_{k_0}) \\ &\leq \tfrac{2}{\gamma_{k_1}}(\delta_{k_1} + f(x^{k_0}) - \bar{f}_{k_0}) \\ &\leq \tfrac{2}{\gamma_{k_1}}(\delta_{k_1} + \tilde{\varepsilon}_{k_0}),\end{aligned}$$

où la seconde inégalité vient de (6.18), des hypothèses sur γ_k et de la décroissance de $\{\delta_k\}$ (pour k suffisamment grand). La dernière inégalité est une conséquence de (6.9). Par conséquent, $\{z^k\}$ est bornée.
Puisque pour $k \geq k_0$, le test de diminution (6.10) n'est jamais satisfait, nous avons que

$$\bar{f}_k - \tilde{f}_{k+1} - \tilde{\varepsilon}_{k+1} < \sigma\delta_k.$$

D'où,

$$\begin{aligned}(1-\sigma)\delta_k &< \tilde{f}_{k+1} + \tilde{\varepsilon}_{k+1} - \bar{f}_k + \delta_k \\ &\leq f(z^{k+1}) + \tilde{\varepsilon}_{k+1} - \varphi_k(z^{k+1}) - \tfrac{\gamma_k}{2}\|z^{k+1} - x^{k_0}\|^2 \\ &\leq f(z^{k+1}) - \tilde{f}_k + \varphi_k(z^k) - \varphi_k(z^{k+1}) + \tilde{\varepsilon}_{k+1} \\ &\leq f(z^{k+1}) - f(z^k) + \tilde{\varepsilon}_k + \varphi_k(z^k) - \varphi_k(z^{k+1}) + \tilde{\varepsilon}_{k+1} \\ &\leq 2L\|z^{k+1} - z^k\| + 2\max\{\tilde{\varepsilon}_i \mid i \in I_k^c\},\end{aligned}$$

où la seconde inégalité vient de (6.9) et (6.7), la troisième vient de $\varphi_k(z^k) \geq l_k(z^k) = \tilde{f}_k$, la quatrième vient de (6.9) et la dernière est une conséquence

de la Lipschitz continuité de f et φ_k (la suite $\{z^k\}$ étant bornée).
En combinant la dernière relation avec (6.21), nous obtenons que

$$\delta_k - \delta_{k+1} \geq \frac{\gamma_k}{8L^2}((1-\sigma)\delta_{k+1} - 2\max\{\tilde{\varepsilon}_i \mid i \in I^c_{k+1}\})^2.$$

Puisque $\delta_k - \delta_{k+1} \to 0$ (par (6.22)) et $\gamma_k \geq \gamma_{k_1} > 0$, cette dernière relation implique que

$$\lim_{k\to\infty} \delta_k = \bar{\delta} = \frac{2}{1-\sigma} \lim_{k\to\infty} (\max\{\tilde{\varepsilon}_i \mid i \in I^c_k\}).$$

En passant à la limite dans (6.20), nous obtenons donc que

$$(2 + \frac{2}{1-\sigma}) \lim_{k\to\infty} (\max\{\tilde{\varepsilon}_i \mid i \in I^c_k\}) \geq \Delta_{\text{opt}},$$

ce qui contredit (6.4).

□

Pour terminer, notons que cette analyse fournit un résultat de stabilité pour la méthode faisceau. En effet, au lieu de s'interroger au sujet des erreurs d'approximation $\tilde{\varepsilon}$ afin de garantir la terminaison de l'algorithme par un point satisfaisant un niveau d'optimalité donné Δ_{opt}, nous pouvons nous poser la question suivante : étant donnée une borne supérieure ϵ sur les erreurs d'approximation $\tilde{\varepsilon}$ i.e.

$$\epsilon \geq \limsup_k(\max\{\tilde{\varepsilon}_i \mid i \in I^c_k\}),$$

Quelles sont les propriétés de convergence de la méthode faisceau ? Quel type de solutions approximées peut-on obtenir ? Comment dépendent-elles de la valeur de ϵ ?
Nos résultats nous garantissent de trouver une solution approximée de (6.1) au sens de (6.2) et (6.3) où

$$\Delta_{\text{opt}} = \frac{2(2-\sigma)\epsilon}{1-\sigma} + t, \quad t > 0 \text{ arbitrairement petit.}$$

6.3 Application : Relaxation Lagrangienne

Considérons le problème primal suivant :

$$\begin{array}{ll} \max_x & q(x) \\ \text{s.c.} & h(x) = 0 \\ & x \in P, \end{array} \tag{6.23}$$

où $P \subset \mathbb{R}^m$ compact, $q : \mathbb{R}^m \to \mathbb{R}$ et $h : \mathbb{R}^m \to \mathbb{R}^n$. La fonction lagrangienne est $L(x, \mu) = q(x) + < \mu, h(x) >$ où $\mu \in \mathbb{R}^n$ et la fonction duale est par conséquent

$$d(\mu) = \max_{x \in P} \; q(x) + < \mu, h(x) > . \tag{6.24}$$

On peut alors écrire le problème dual

$$\min_{\mu \in \mathbb{R}^n} \; \max_{x \in P} \; q(x) + < \mu, h(x) > . \tag{6.25}$$

Dans certaines situations, il est intéressant d'essayer de résoudre le problème primal (6.23) en résolvant le problème dual (6.25). La relaxation lagrangienne donne lieu au problème convexe non nécessairement différentiable (6.25) indépendamment de toutes hypothèses au sujet de (6.23). Bien entendu, sans aucune autre hypothèse sur (6.23), il peut exister un saut de dualité entre les valeurs optimales primale et duale. Par ailleurs, après avoir résolu le problème dual, il faut encore retrouver les solutions primales. Malgré cela, cette approche reste intéressante pour certaines applications puisqu'elle fournit une borne supérieure sur la valeur optimale de (6.23).
Puisque la fonction duale est convexe et non nécessairement différentiable, la méthode faisceau du chapitre 3 peut être utilisée pour résoudre (6.25). Rappelons que la méthode faisceau reposait sur l'existence présumée d'un oracle (Hypothèse 3.0.1). La proposition suivante indique en quoi l'oracle consiste réellement pour la fonction duale (6.24).

Proposition 6.3.1 *Soit $\mu \in \mathbb{R}^n$. Si x^* est solution de (6.24), alors*

1. $d(\mu) = q(x^*) + < \mu, h(x^*) >$,
2. $h(x^*) \in \partial d(\mu)$.

Preuve :

1. Evident.
2. En effet, $\forall y \in \mathbb{R}^n$:

$$\begin{aligned} d(y) &= \max_{x \in P} q(x) + < y, h(x) > \\ &\geq q(x^*) + < y, h(x^*) > - < u, h(x^*) > + < u, h(x^*) > \\ &= d(u) + < h(x^*), y - u > . \end{aligned}$$

□

Autrement dit, supposer l'existence d'un oracle pour la fonction duale revient à supposer que l'on peut calculer exactement la solution de (6.24) pour tout $\mu \in \mathbb{R}^n$. Or, calculer la solution exacte de ce problème peut être très coûteux et numériquement irréalisable puisque les méthodes utilisées pour ce type de problème renvoient une solution approximée en accord avec un critère (fini) de terminaison. Par ailleurs, il vaut mieux ne pas passer plus de temps que de nécessaire pour résoudre ce problème puisque ce n'est pas l'objectif premier. C'est dans cette optique que la méthode de Solodov est d'un grand intérêt. En effet, les hypothèses de Solodov nécessitent que l'on sache calculer $\tilde{d} \in \mathbb{R}$ et $y \in \mathbb{R}^n$ tels que

$$d(\mu) \geq \tilde{d} \geq d(\mu) - \tilde{\varepsilon},$$
$$d(z) \geq \tilde{d} + < y, z - \mu > \quad \forall z \in \mathbb{R}^n,$$

pour tout $\mu \in \mathbb{R}^n$. La proposition suivante indique que ces hypothèses seront satisfaites en calculant un $\tilde{\varepsilon}$-maximum du problème (6.24).

Proposition 6.3.2 *Soit $\mu \in \mathbb{R}^n$. Si x^* est un ε-maximum de (6.24), alors $\tilde{d} = q(x^*) + < \mu, h(x^*) >$ et $y = h(x^*)$ vérifient les hypothèses de Solodov.*

Preuve :

1. $d(\mu) \geq q(x^*) + < \mu, h(x^*) > \geq d(\mu) - \tilde{\varepsilon}$ par définition de x^*.
2. $\forall z \in \mathbb{R}^n$,

$$\begin{aligned} d(z) &= \max_{x \in P} q(x) + < z, h(x) > \\ &\geq q(x^*) + < z, h(x^*) > - < \mu, h(x^*) > + < \mu, h(x^*) > \\ &= q(x^*) + < \mu, h(x^*) > + < h(x^*), z - \mu > . \end{aligned}$$

□

Par conséquent, dans le contexte de la relaxation lagrangienne, la méthode de Solodov se contente du calcul d'un $\tilde{\varepsilon}$-maximum de (6.24) à chaque itération.

6.3.1 Résultats numériques

Considérons le problème quadratique suivant

$$\begin{array}{ll} \max & \frac{1}{2}x^T Hx + f^T x \\ \text{s.c.} & Ax = b \\ & x \in [-10, 10]^4 \end{array} \tag{6.26}$$

où $H = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -0.5 & 1 \\ 0 & 0 & 1 & -7 \end{pmatrix}$, $f = \begin{pmatrix} -1 \\ 0 \\ 2 \\ -7 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 5 & 2 & -3 \\ 2 & 0 & 1 & -1 \\ 3 & -2 & 5 & 0 \\ 4 & 0 & 2 & -2 \end{pmatrix}$

et $b = \begin{pmatrix} -2 \\ 3 \\ 4 \\ 6 \end{pmatrix}$.

On se propose d'approcher la valeur optimale du problème quadratique (6.26) via la méthode de Solodov appliquée à son dual. Autrement dit, nous allons minimiser la fonction duale

$$d(y) = \max_{x \in [-10,10]^4} \frac{1}{2}x^T Hx + f^T x + y^T(Ax - b).$$

L'implémentation a été réalisée en langage MATLAB. Les principales caractéristiques de l'algorithme sont :

. le paramètre σ est initialisé à 0.4,
. le point de départ est $y_0 = (1, 1, 1, 1)$,
. le paramètre γ_k est constant pendant tout le processus et vaut 10.

Code MATLAB :

```
fip=fopen('Résultat.doc','w');

H= zeros(4,4);
H(1,1) = -1;H(2,2)=-2;H(3,3)=-1/2;H(4,4)=-7;H(4,3)=1;H(3,4)=1;

f=[-1;0;2;-7];
```

```
A=[1 5 2 -3;2 0 1 -1;3 -2 5 0;4 0 2 -2];
b=[-2;3;4;6];

LB(1)=-10;LB(2)=-10;LB(3)=-10;LB(4)=-10;
LU(1)=10;LU(2)=10;LU(3)=10;LU(4)=10;

fprintf(fip,'Problème de départ : \n');
fprintf(fip,'maximiser 1/2*x*H*x+f*x \n');
fprintf(fip,'s.c. A*x=b où \n');
fprintf(fip,'\n');
fprintf(fip,'H= \n');
for i=1:4
    fprintf(fip, '%3.1f   %3.1f   %3.1f   %3.1f \n', H(i,1),
                                    H(i,2), H(i,3), H(i,4));
end
fprintf(fip,'\n');
fprintf(fip,'A= \n');
for i=1:4
    fprintf(fip, '%3.1f   %3.1f   %3.1f   %3.1f \n', A(i,1),
                                    A(i,2), A(i,3), A(i,4));
end
fprintf(fip,'\n');
fprintf(fip,'f= \n');
for i=1:4
    fprintf(fip, '%3.1f \n', f(i));
end
fprintf(fip,'\n');
fprintf(fip,'b= \n');
for i=1:4
    fprintf(fip, '%3.1f \n', b(i));
end
fprintf(fip,'\n');
fprintf(fip,'-10<=x(i)<=10 i=1,...,4 \n');
fprintf(fip,'\n');

a0=[rand(1);rand(1);rand(1);rand(1)];

[a,FVAL,EXITFLAG,OUTPUT]=QUADPROG(-H,-f,[],[],A,b,LB,LU,a0);
FVAL=-FVAL
```

```
fprintf(fip,'Valeur optimale du primal :%4.6f \n',FVAL);
fprintf(fip,'\n');
fprintf(fip,'\n');

% Algorithm de Solodov

% initialisation

t=cputime;
sigma=0.4;
gamma=10;
G= zeros(5,5);
for j=1:4
    G(j,j) = gamma;
end

% Tolérance

epsilon=0.0001;
deltaopt=0.00055;

%TolFun - Termination tolerance on the function value
%                                    [ positive scalar ]

OPTIONS = OPTIMSET('Tolfun',epsilon);

% Dimension du faisceau

Bmax=5;
Bcompteur=0;
B=[];
ineq=[];

% Point de départ

x=[1;1;1;1];

fprintf(fip,'Algorithme de Solodov pour la relaxation
                                      lagrangienne \n');
fprintf(fip,' \n');
```

```
fprintf(fip,'sigma=%1.4f \n',sigma);
fprintf(fip,'gamma=%3.2f \n',gamma);
fprintf(fip,'epsilon=%3.5f \n',epsilon);
fprintf(fip,'Delta optimal=%3.5f \n',deltaopt);
fprintf(fip,'Taille maximale du faisceau=%3.0f\n',Bmax);
fprintf(fip,'Point de départ= %3.5f  %3.5f  %3.5f  %3.5f \n',
                                        x(1),x(2),x(3),x(4));
fprintf(fip,'\n');
fprintf(fip,' Iter Fappr norm(sousgrad) \n');
fprintf(fip,'\n');
fbis=f+A'*x;
[a,FVAL,EXITFLAG,OUTPUT,LAMBDA]=QUADPROG(-H,-fbis,[],[],[],[],
                                        LB,LU,a0,OPTIONS);
Sousgrad=A*a-b;
Fappr=-FVAL-x'*b
k=0;
fprintf(fip,' %3.0f \t %5.4f \t %5.4f \n',k, Fappr,
                                    norm(Sousgrad));
z=x;
fbar=Fappr;
lbar1max= [Sousgrad' -1];
lbar2max= [Sousgrad'*z-Fappr];
varquad = [rand(1);rand(1);rand(1);rand(1);rand(1)];
for k=1:100
    fobj=[-gamma*x; 1];
    lbar1= [Sousgrad' -1];
    lbar2= [Sousgrad'*z-Fappr];
    B= [B; lbar1]; ineq=[ineq;lbar2];
    Bcompteur=Bcompteur+1;
    varquad = quadprog(G,fobj,B,ineq,[],[],[],[],varquad);
    z=varquad;
    z(5)=[];
    model=varquad(5);
    delta=fbar-model-gamma/2*norm(z-x)^2;
    if delta+2*epsilon <= deltaopt
        break
    end
    fbis=f+A'*z;
    [a,FVAL,EXITFLAG,OUTPUT,LAMBDA]=QUADPROG(-H,-fbis,[],[],
                                    [],[],LB,LU,a,OPTIONS);
```

```
    Sousgrad=A*a-b;
    Fappr=-FVAL-z'*b
    if fbar-Fappr-epsilon >= sigma*delta
        lbar1max= lbar1;
        lbar2max= lbar2;
        if Bmax <= Bcompteur
            B=[];
            B=[B;gamma*(x-z)' -1;lbar1max];
            ineq=[];
            ineq=[ineq;gamma*(x-z)'*x-gamma*norm(z-x)^2-model;
                                                    lbar2max];
            Bcompteur=2;
        end
        x=z;
        fbar=Fappr;
        fprintf(fip,' %3.0f \t %5.6f \t %5.4f  \n', k, Fappr,
                                              norm(Sousgrad));
    else
        if Bmax <= Bcompteur
            B=[];
            B=[B;gamma*(x-z)' -1;lbar1max];
            ineq=[];
            ineq=[ineq;gamma*(x-z)'*x-gamma*norm(z-x)^2-model;
                                                    lbar2max];
            Bcompteur=2;
        end
    end
end
t=cputime-t;
fprintf(fip,'\n');
fprintf(fip,'\n');
fprintf(fip,'Valeur optimale duale finale=%4.6f \n',Fappr);
fprintf(fip,'norme du sous-gradient final=%4.6f \n',
                                      norm(Sousgrad));
fprintf(fip,'\n');
fprintf(fip,'cputime (in seconds) : %8.4f',t);
fclose(fip);
```

Notons que la valeur optimale de (6.26) est 0.085577.
L'algorithme avec $B_{\max} = 5$ a été lancé avec différents niveaux d'optimalité Δ_{opt} afin de mesurer l'influence de ce paramètre sur le comportement général de la méthode. Notons que les degrés de précision $\tilde{\varepsilon}$ (constants pendant le processus d'optimisation) ont été choisis de manière à vérifier la relation (6.4). Dans le tableau suivant, k correspond au nombre d'itérations et d^* correspond à la valeur optimale (du dual) obtenue.

Δ_{opt}	$\tilde{\varepsilon}$	k	d^*
0.55	0.1	5	0.138289
0.055	0.01	10	0.091966
0.0055	0.001	20	0.085736
0.00055	0.0001	17	0.085587

Les résultats de ce tableau se résument de la façon suivante :

$$\text{si } \Delta_{\text{opt}} \searrow \text{ alors } \tilde{\varepsilon} \searrow, \ k \nearrow \text{ et } d^* \searrow 0.085577.$$

Autrement dit, plus le niveau d'optimalité désiré est grand, plus le paramètre $\tilde{\varepsilon}$ doit être faible (pour assurer la terminaison), plus le nombre d'itérations est grand et plus la valeur optimale obtenue s'approche de la valeur optimale du primal.
Dans le tableau suivant, CPU correspond à un indice du temps d'exécution (CPU=1 pour $B_{\max} = 10$).

$B_{\max}$	k	CPU
2	39	6.616
5	17	1.001
10	15	1

Selon ce tableau, plus la taille maximale du faisceau est grande et plus le nombre d'itérations et le temps d'exécution sont faibles.

Chapitre 7

Conclusion

Les différentes méthodes développées dans ce mémoire ont été construites en appauvrissant successivement les hypothèses de départ. Les hypothèses les plus faibles étant celles de la méthode de Solodov, on remarque malgré tout que celle-ci reste simple, efficace et paramètrable à souhait. D'autre part, cette méthode repose sur une théorie précise et riche en informations commentant la façon dont doit être implémentée la méthode en fonction du degré d'optimalité souhaité.

Pour terminer, notons qu'aucun résultat numérique n'a été fourni pour la méthode du point proximal inexact et ce, pour une raison très simple. La théorie corrigée du chapitre 5 semble correcte. Cependant, les hypothèses imposées sur le modèle ne paraissent pas à première vue être simples à satisfaire. Par ailleurs, les conditions (5.16)-(5.19) imposées sur les tolérances et sur les approximations de la fonction ne semblent pas plus évidentes à vérifier. Il serait donc intéressant de vérifier le bien fondé et la consistance pratique de toutes ces hypothèses.

Bibliographie

[1] M. Florenzano C. Le Van, *Finite dimensional convexity and optimization*, Springer-Verlag, 2001.

[2] Antonio Frangioni, *Generalized Bundle Methods*, SIAM, 2002.

[3] C. Helmberg F. Rendl, *A Spectral Bundle Method for Semidefinite Programming*, SIAM, 2000.

[4] Jean-Baptiste Hiriart-Urruty et Claude Lemaréchal, *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, 1993.

[5] Jean-Baptiste Hiriart-Urruty et Claude Lemaréchal, *Convex Analysis and Minimization Algorithms II*, Springer-Verlag, 1993.

[6] Scott A. Miller, *Thèse : An inexact bundle method for solving large structured linear matrix inequalities*, décembre 2001.

[7] M. V. Solodov, *On approximations with finite precision in bundle methods for nonsmooth optimization*, Journal of Optimization Theory and Applications, 2002.

[8] Jean-Jacques Strodiot, *Notes de cours : Nonsmooth optimization Theory and Algorithms.*